

2013

Exploring and understanding the substrate diversity of β -ketoacyl-ACP synthase III enzymes

Shivani Garg
Iowa State University

Follow this and additional works at: <https://lib.dr.iastate.edu/etd>

 Part of the [Biochemistry Commons](#)

Recommended Citation

Garg, Shivani, "Exploring and understanding the substrate diversity of β -ketoacyl-ACP synthase III enzymes" (2013). *Graduate Theses and Dissertations*. 16085.
<https://lib.dr.iastate.edu/etd/16085>

This Dissertation is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Graduate Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.

**Exploring and understanding the substrate diversity of
 β -ketoacyl-ACP synthase III enzymes**

by

Shivani Garg

A dissertation submitted to the graduate faculty
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Molecular, Cellular and Developmental Biology

Program of Study Committee:
Guru Rao, Co-Major Professor
Drena Dobbs, Co-Major Professor
Basil J. Nikolau
Marna D. Yandea-Nelson
Thomas Bobik
Edward Yu

Iowa State University

Ames, Iowa

2013

Copyright © Shivani Garg, 2013. All rights reserved.

*“Two great things you can give your children:
one is roots, the other is wings.” – Hodding Carter*

Dedicated to my grandfather and parents who gave me both roots and wings

TABLE OF CONTENTS

	Page
DEDICATION	ii
ACKNOWLEDGEMENTS	v
ABSTRACT	vii
CHAPTER I. GENERAL INTRODUCTION.....	1
Introduction to β -ketoacyl-ACP synthase III.....	1
Questions explored in the dissertation	4
Dissertation organization	5
Literature review	7
Significance	16
References	17
 CHAPTER II. DISSECTING THE STRUCTURAL BASIS OF DIVERSE FUNCTIONALITIES OF <i>E. COLI</i> AND <i>B. SUBTILIS</i> KASIII ENZYMES USING STD NMR BASED LIGAND INTERACTIONS	 23
Abstract	23
Introduction	24
Materials and methods	28
Results	33
Discussion	41
Tables	47
Figures	49
References	55
 CHAPTER III. DELINEATING THE STRUCTURE-FUNCTION RELATIONSHIPS OF β-KETOACYL-ACP SYNTHASE III BASED ON PHYLOGENETIC AND FUNCTIONAL COMPARISONS	 59
Abstract	59
Introduction	60
Materials and methods	63
Results	66
Discussion	73

	Page
Tables	81
Figures	83
References	90
 CHAPTER IV. IDENTIFICATION OF KASIII ENZYMES WITH NOVEL SUBSTRATE SPECIFICITIES: DEMONSTRATION OF <i>IN VIVO</i> PRODUCTION OF NOVEL ω-1 HYDROXYLATED FATTY ACIDS USING A NOVEL KASIII	
Abstract	94
Introduction	95
Materials and methods	97
Results	104
Discussion	110
Tables	116
Figures	118
References	123
 CHAPTER V. CONCLUSIONS.....	
Conclusions.....	129
References	139
 APPENDIX. BUSINESS PLAN FOR OMEGACHEA BIORENEWABLES LLC	
	142

ACKNOWLEDGEMENTS

I would like to thank my advisor, Dr. Basil J. Nikolau, who gave me the opportunity to work on this exciting research topic. This work would have been impossible without his constant support and guidance throughout my graduate school years. I am so grateful for the numerous opportunities and enriching experiences that I had in his lab such as mentoring REU students, developing collaborations, undertaking biorenewable chemicals minor and exploring commercial aspects of my research; these experiences along with his confidence in my abilities have helped me grow tremendously as a researcher and as a person.

Next, I would like to thank Dr. Marna D. Yandeu-Nelson, my co-advisor, who provided invaluable advice and mentorship contributing greatly to the development of this work. I also want to thank my POS committee members, Dr. Thomas Bobik, Dr. Edward Yu, Dr. Drena Dobbs and Dr. Guru Rao for their valuable suggestions and willingness to help at all times.

This research was funded by NSF and CBiRC, and I am grateful for the interdisciplinary research environment of CBiRC that provided unique opportunities to extend my research into areas I had not imagined before. Dr. Peter Keeling, CBiRC Innovation Director, was instrumental in helping me develop a commercialization path for my research. Most cherished are my conversations with him that broadened my vision and enabled me to look at research from an industrial point of view. The constant encouragement from Dr. Brent Shanks, CBiRC Director, motivated me to do better at my research. I am thankful for the collaboration opportunities that CBiRC provided, including a research visit to Dr. Joseph Noel's lab at the Salk Institute for Biological Studies, which markedly

enhanced my research and technical skills. I am thankful to Dr. Charles Stewart, research specialist in the Noel lab, for his expert mentorship and for teaching me thermal shift assays and crystallization techniques that brought a new perspective to my research.

I want to thank Dr. Bruce Fulton, NMR facility manager, who facilitated saturation transfer difference NMR experiments and helped me analyze the NMR experimental results. I would like to acknowledge the W. M. Keck Metabolomics lab where enzyme activity assays were conducted, and Dr. Ann Perera was helpful in solving technical issues related to instrumentation. I highly appreciate the time and the sincerest efforts that my undergraduate students, Luke Prest, Kim-Xuan Nguyen, Morgan Becker, Joshua Scaralia, Ruben Garcia and Niibari Menegbo, devoted to this research.

Amy Arhens and Katie Blair, former and current MCDB program coordinators, were always accessible to answer my questions regarding administrative and program completion requirements. I deeply appreciate their timely help and advice whenever I needed it.

I wish to thank all the members of the Nikolau lab, especially Dr. Libuse Brachova, for continuous support and constructive criticism that immensely enhanced the quality of this work. My great room-mates and friends, both in and out of Ames, kept me a good company and supported me through the ups and downs of my graduate school years.

I am deeply indebted to my parents and late grand-parents. Without their unconditional love, encouragement and endless support, I would not have realized this dream. Equally indebted I am to my brother, Nishant Garg, who has been a great source of inspiration and always had my back in all my endeavors including this one.

ABSTRACT

Fatty acids are an integral part of the lipid membrane of bacteria and plants, and the enzyme that initiates fatty acid synthesis, β -ketoacyl-ACP Synthase III (KASIII), is one of the most diverse enzymes found in nature. KASIII exhibits vast diversity in its substrate specificity. By virtue of its substrate specificity, KASIII determines the fatty acid profile of the organism by dictating the nature of the ω -end of the fatty acids. For example, KASIIIs from many bacteria (e.g. *Escherichia coli*) accept straight chain acyl-CoAs resulting in production of straight chain fatty acids and are, therefore, narrow in substrate specificity. In contrast, KASIII from other bacteria (e.g. *Staphylococcus aureus*) accept both straight and branched chain acyl-CoAs resulting in production of both straight and branched chain fatty acids and therefore, exhibit broad substrate specificity. Despite the availability of KASIII crystal structures from a dozen sources, the molecular basis for its substrate diversity remains unclear. Additionally, the enzymatic activity of KASIII enzymes has not been explored with substrates other than straight or branched chain acyl-CoAs. However, occurrence of ω -cyclic and ω -phenylic fatty acids in certain bacteria and plants suggests that KASIII in such organisms may be capable of utilizing atypical substrates such as cyclic acyl-CoAs or aromatic acyl-CoAs.

In this dissertation, we have experimentally established the essentiality of three residues in determining KASIII substrate binding and catalysis. We have also characterized twelve putative KASIII enzymes from diverse bacterial sources and have identified six functional KASIIIs from this set. Additionally, based on phylogenetic and functional comparisons of all the characterized KASIIIs, we have developed a structure-function

correlation that can be used for predicting KASIII functionality. Finally, we identified three novel KASIII enzymes that can utilize atypical substrates, including hydroxylated, aromatic and unsaturated acyl-CoAs. We subsequently demonstrated the *in vivo* production of novel ω -1-hydroxy-branched fatty acids using a novel KASIII. The knowledge gained from this study has enhanced our understanding of the KASIII structure-function relationship and has paved the way for developing fatty acids with novel chemical functionalities at the ω -end using KASIII with diverse specificities.

CHAPTER I

GENERAL INTRODUCTION

Introduction to β -Ketoacyl-ACP Synthase III

Biosynthesis of fatty acids is an essential metabolic process for the generation of phospholipids and cellular membranes. Mammalian fatty acid synthesis is catalyzed via a type I fatty acid synthase that comprises a multi-modular enzyme complex, which carries out iterative steps of the fatty acid synthesis (FAS) cycle. In contrast, in most bacteria and plants, fatty acid synthesis proceeds via a type II system that has discrete enzymes for catalyzing each step of FAS. The enzyme that initiates as well as acts as a 'pacemaker' (i.e. the rate limiting step) of type II FAS is β -ketoacyl-acyl carrier protein synthase III (KASIII or FabH) (1,2). It initiates *de novo* fatty acid synthesis via a Claisen-type condensation of an acyl-coenzyme A (or acyl-CoA) molecule with malonyl-acyl carrier protein (malonyl-ACP) to produce a β -ketoacyl-ACP (*figure*

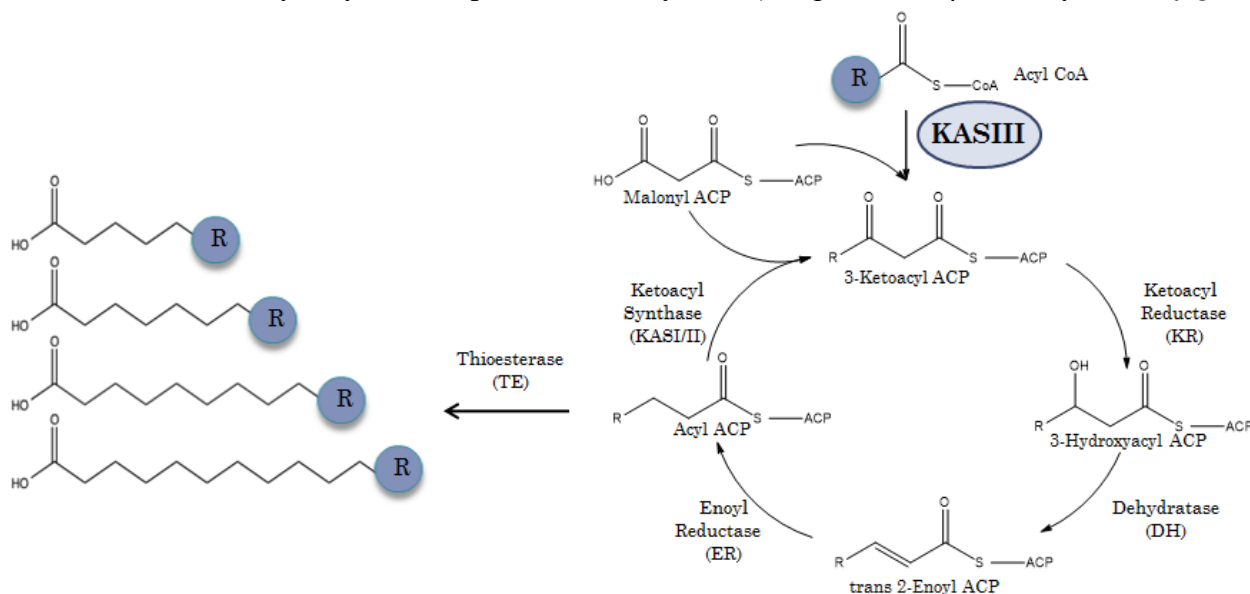


Figure 1. KASIII initiates the type II fatty acid synthesis (FAS) cycle in bacteria and plants. KASIII catalyzes Claisen-like condensation of a starter molecule, an acyl CoA, with an extender molecule, typically malonyl ACP, to give a ketoacyl-ACP intermediate that undergoes successive steps of reduction, dehydration and reduction to give a reduced acyl ACP.

1) intermediate, that undergoes iterative fatty acid synthesis steps of reduction, dehydration, reduction and elongation until a saturated fatty acyl-ACP of typically 14-18 carbons is formed (3).

Depending on the nature of the R group in acyl CoA, KASIII is capable of introducing various functionalities into the terminal ω -end (omega end) of the fatty acid molecule (*figure 2*). Therefore, besides initiating fatty acid synthesis, KASIII also determines the fatty acid profile of an organism by virtue of its substrate specificity. For example, KASIIIs from Gram-negative bacteria such as *Escherichia coli* have narrow substrate specificities with preferences for short straight chain acyl-CoA primers (acetyl-CoA, propionyl-CoA) (4-8), whereas KASIIIs from Gram-positive bacteria, such as *Bacillus subtilis* (9) or *Staphylococcus aureus* (10), display broad substrate specificities and can use both straight and branched chain acyl-CoAs (e.g. isobutyryl-CoA, isovaleryl-CoA), resulting in formation of both straight and branched chain fatty acids. In addition to the utilization of straight and branched chain primers, KASIII in several bacteria has been shown to utilize long chain acyl primers. For example, the KASIII from *Mycobacterium tuberculosis* utilizes lauryl CoA resulting in formation of mycolic acids (11).

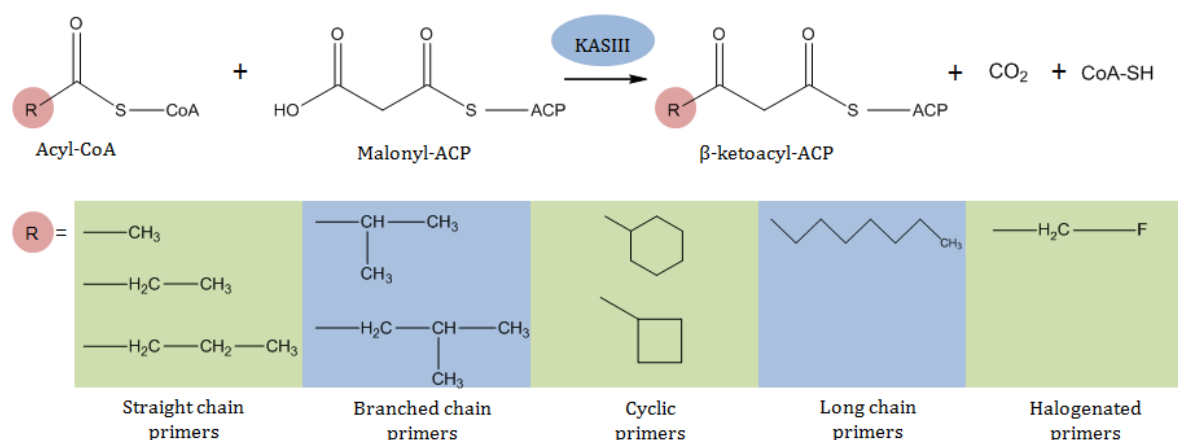


Figure 2. KASIII determines the fatty acid profile of an organism. The 'R' group of the starter molecule, acyl CoA, could be straight-chain, branched-chain, cyclic, long-chain or halogenated, resulting in different types of fatty acid products.

KASIII has been most extensively studied in *E. coli*, and has been crystallized both in the presence and absence of ligands (12-14). KASIII crystal structures are available from eight other bacterial sources (15-21) and attempts have been made to understand the underlying structural basis of their substrate promiscuities (17). However, it yet remains to be understood which specific amino acid residues or structural motifs govern KASIII substrate preferences. It has been hypothesized that the rotamer conformation of a particular conserved Phe residue (corresponding to Phe304 in *E. coli* KASIII) may be responsible for differences in narrow and broad substrate specificity KASIII enzymes (17,18). This hypothesis has not yet been experimentally verified, but evidence from sequence alignments of numerous KASIIIs corroborates this theory (17).

Thus far, KASIII substrate specificity has been explored with only three classes of acyl-CoA primers – a) straight chain primers, such as acetyl-CoA, propionyl-CoA, butyryl- or hexanoyl-CoA (5,22-24), b) branched chain primers, such as isobutyryl-CoA or isovaleryl-CoA (10,22,23), and c) long chain straight primers, such as lauryl-CoA (25). Given the vast diversity in KASIII enzymes, it may be possible that KASIII enzymes are capable of utilizing substrates other than the three substrate classes listed above. In fact, the presence of unique fatty acids in some organisms, such as fatty acids with cyclic groups at their ω -ends in *Alicyclobacillus acidocaldarius* (26), or fatty acids with fluorinated groups at their terminal ω -ends in *Dichapetalum toxicarium* (27), suggests that KASIII in such organisms is capable of utilizing unique primers such as cyclic acyl CoA or fluorinated acyl CoAs, thereby incorporating these unique functionalities into the ω -end of final fatty acid products. Such KASIIIs have not yet been characterized, but these examples illustrate that KASIII enzymes may be capable of condensing unique acyl CoA substrates (*figure 2*), such as cyclic acyl CoA, hydroxylated acyl CoA, aromatic

or even unsaturated acyl CoAs, resulting in novel ω -functionalized fatty acids. This concept presents a new direction for research on KASIII enzymes.

Questions being explored in the dissertation

Given the wide scope of functional divergence displayed by KASIII enzymes, the factors that govern KASIII functional promiscuity remain unanswered. Despite the availability of several KASIII crystal structures (10,12,28-30) and proposed hypotheses regarding several motifs and residues that may affect substrate specificity (17), there is no definite and clear understanding of what determines KASIII substrate specificity. Additionally, although KASIII enzymes are known to utilize a variety of substrates, specifically straight chain, branched chain and long straight chain acyl-CoAs (*figure 3*), their functionalities with many other acyl-CoA substrates (e.g. hydroxylated, cyclic, aromatic or unsaturated acyl-CoAs) have not yet been explored. Further investigation in these areas will bridge the gap between KASIII structural and

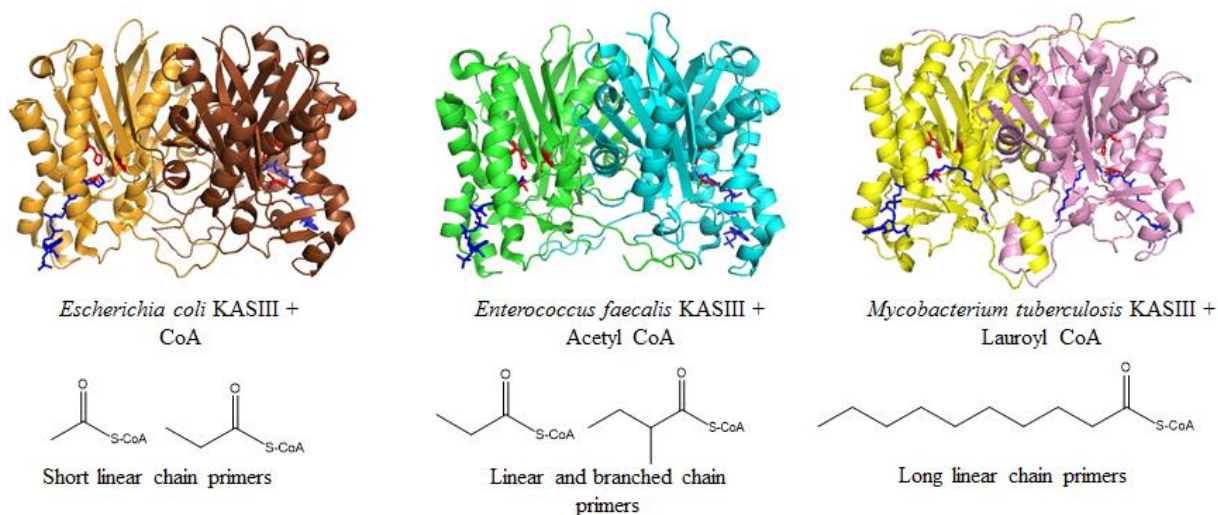


Figure 3. Structural and functional diversity in KASIII enzymes. Co-crystal structures of KASIII from *E. coli* (PDB code – HND), *E. faecalis* (PDB code – 3IL4), and *M. tuberculosis* (PDB code – 1U6S) are depicted with some of their preferred substrates.

functional information, and will enable development of novel KASIII biocatalysts that could be used for synthesizing unique ω -functionalized fatty acids with applications in the bio-based chemical industry.

The goal of this dissertation is two-fold. Firstly, this body of work aims at bridging the knowledge gap between the known functional diversity of KASIII enzymes and the structural basis for that diversity, and identifying the underlying molecular basis for KASIII substrate diversity. Secondly, this research aims to characterize and understand the vast and largely unexplored biological diversity of KASIII enzymes, so that ultimately this diversity can be exploited to produce novel fatty acids with new functionalities at their ω -ends (e.g. hydroxyl, aromatic or cyclic groups).

Dissertation organization

This dissertation investigates the structure-function relationships of KASIII enzymes, and also explores their great functional diversity to identify unique KASIII biocatalysts that can be used to synthesize novel fatty acids.

The dissertation begins in Chapter 1 with a general introduction to KASIII enzymes followed by a comprehensive survey of published literature related to KASIII enzyme structure and function.

Chapter 2 presents the determination of specific amino acid residues that govern substrate specificity of KASIII enzymes. Biochemical and biophysical approaches (*in vitro* spectrophotometric enzyme activity assay and saturation transfer difference NMR) are applied to KASIII enzymes that exhibit narrow (*E. coli* KASIII) or broad (*B. subtilis* KASIIIa and

KASIIIb) substrate specificities. The experiments were designed and executed by Shivani Garg under the supervision of Marna Yandea-Nelson and Basil Nikolau.

Chapter 3 builds upon the knowledge gained within chapter 2 and demonstrates how this knowledge can be used to predict the functionality of KASIII enzymes. The functionalities of ten previously uncharacterized, putative KASIII enzymes are first predicted based on learning from previous work described in chapter 1 and our experimental results from chapter 2, and are then tested using biochemical and biophysical approaches (*in vitro* enzyme activity assays and thermal shift assays). Within this study, we also determine functional divergence or redundancy among KASIII homologs derived from the same bacterial strains. The work in this chapter was collaboration between Shivani Garg and Charles Stewart (The Salk Institute for Biological Studies). This work was conducted under the supervision of Marna Yandea-Nelson and Basil Nikolau at ISU, and Joseph Noel at Salk Institute for Biological Studies. Shivani Garg and Charles Stewart designed and executed the thermal shift assay experiments at the Salk Institute for Biological Studies, and Shivani Garg continued with activity assays and sequence analysis of selected KASIII genes at Iowa State University. The journal chapter was prepared by Shivani Garg, under the guidance of Marna Yandea-Nelson and Basil Nikolau.

After investigating structure-function relationships and exploring the functional diversity of KASIII enzymes, Chapter 4, focuses on identifying KASIII enzymes with novel substrate specificities. We report an *in vivo* and an *in vitro* screen to identify KASIII enzymes that are capable of utilizing unusual substrates such as hydroxylated, aromatic or unsaturated acyl-CoAs. Discovery of such novel KASIII enzymes paves the way for synthesizing novel functionalized fatty acids that can have potential applications in the chemical industry. The work described in this chapter was collaboration between Shivani Garg, Huanan Jin (Iowa State University) and

Charles Stewart (The Salk Institute for Biological Studies) under the supervision of Marna Yandeau-Nelson, Joseph Noel and Basil Nikolau. Experiments were conceived by Shivani Garg under the guidance of Marna Yandeau-Nelson and Basil Nikolau. Experiments related to *B. subtilis* deletion mutant and *R. rubrum* constructs were executed by Huanan Jin. Shivani Garg and Charles Stewart executed the thermal shift assays for selected KASIII genes at the Salk Institute for Biological Studies. This dissertation chapter was prepared by Shivani Garg, with inputs from Huanan Jin, and under the guidance of Marna Yandeau-Nelson and Basil Nikolau.

Chapter 5 summarizes and synthesizes the new knowledge gained from the research presented in this dissertation and discusses the potential impact and future directions emanating from this body of research.

All the chapters in this dissertation were prepared by Shivani Garg, under the guidance and supervision from Marna Yandeau-Nelson and Basil Nikolau.

Literature review

Presented herein is a survey of existing literature describing KASIII enzymes, which includes the research previously conducted to both structurally and functionally classify and characterize KASIII enzymes from various biological sources. This includes initial work toward elucidating the KASIII reaction mechanism and understanding the role of specific residues in determining substrate specificity.

Classification of KASIII enzymes as thiolases

KASIII (also termed FabH) enzymes belong to the thiolase superfamily of enzymes, which consists of enzymes capable of carrying out Claisen condensation in either a decarboxylative or non-decarboxylative manner (31). KASIII is a decarboxylating condensing

enzyme and it occurs either as a part of the multi-modular type I fatty acid synthase (in mammals, fungi and some bacteria), or as a distinct functional unit in the type II dissociated fatty acid synthesis system (in most bacteria and plants). KASIII differs from the other two KAS enzymes of this class, KASI (or FabB) and KASII (or FabF), as it is specific for only acyl-CoA primers, which KASIII condenses with malonyl-ACP (4,6). KASI and KASII enzymes use acyl-ACP primers that are primed by malonyl-ACP, resulting in elongation of products of KASIII and formation of long chain fatty acids (3). KAS enzymes are closely related to another class comprising highly diverse Polyketide Synthases (PKS) that catalyze synthesis of functionally diverse polyketides (32,33).

Structural and functional characterization of KASIII

KASIII, a homodimer of molecular weight of approximately 70 kDa, was first identified in *E. coli* in the late 1980s due to the presence of residual condensing activity after inhibiting with cerulenin the other two KAS enzymes known at that time (34). Since then, *E. coli* KASIII has been studied extensively; it has been used in drug-design studies as a target for designing new antibacterial drugs (35-38), and has been structurally and functionally characterized. *E. coli* KASIII has been crystallized in the presence and absence of ligand by Davies *et al.*, Qiu *et al.* and Gajiwala *et al.* (12,28,29). In total, 37 crystal structures have been solved from 12 different bacterial species, including *E. coli*, *Mycobacterium tuberculosis* (20,21,30), *Staphylococcus aureus* (10), *Enterococcus faecalis* (29) and several more (11,18,39). KASIII has also been functionally characterized from several other bacterial, plant and protozoan sources (9,39-45).

Mechanism of Claisen condensation catalyzed at KASIII active site

The active site residues and substrate binding pocket are well conserved among KASIIIs from different species. Three residues, Cys112, His244 and Asn274, form the catalytic triad in *E.*

coli KASIII and carry out the Claisen condensation of acyl-CoA and malonyl-ACP through two half reactions via the ping-pong mechanism (3). Qiu *et al.* (12) and Davies *et al.* (28) have proposed different chemical mechanisms for the transacylation reaction at the active site citing different molecules for enhancing the nucleophilicity of the active site Cys112. Summarized herein is the mechanism proposed by Davies *et al.*, which is supported by crystallographic data (a modified representation of the mechanism by Davies *et al.* is shown in **figure 4**) –

First half reaction: Acyl-CoA: enzyme transacylation – The first step is the transfer and covalent linkage of the acyl group from the acyl-CoA primer to the Cys112 residue of KASIII. As per Davies *et al.*, the thiol group of Cys112 is deprotonated by the dipole effect of the α -helix in which it is located. The resulting nucleophilic thiolate ion on Cys112 attacks acyl-CoA and forms a thioester with the acyl group, with the release of CoA-SH.

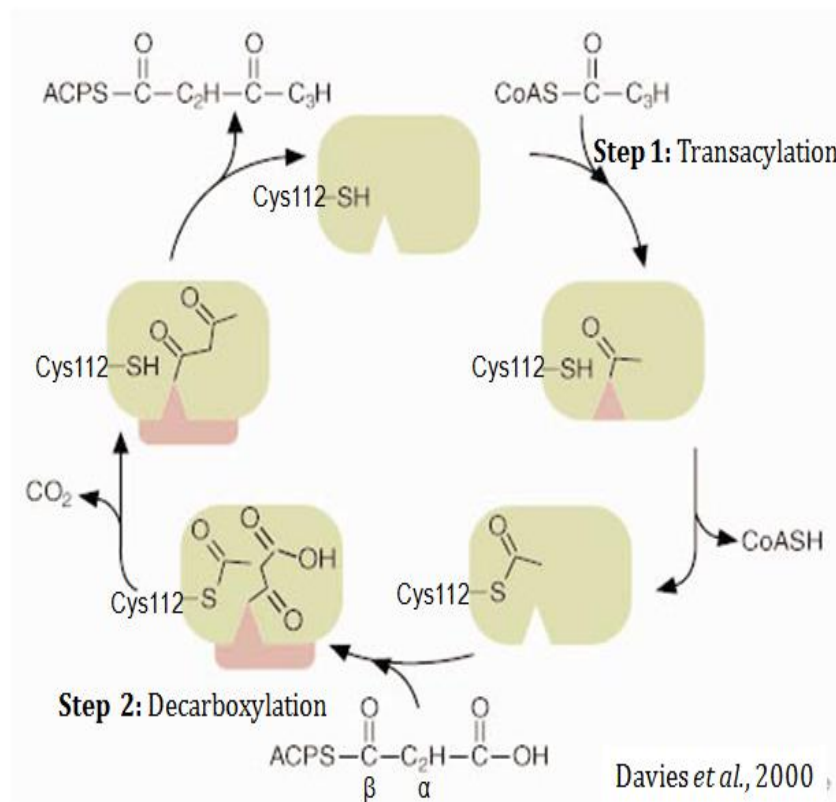


Figure 4. Proposed reaction mechanism of *E. coli* KASIII. In the first step (transacylation), the acetyl group is transferred from acetyl-CoA to the Cys112 of the enzyme, with the concomitant release of CoASH. In the second step (decarboxylation), malonyl-ACP is decarboxylated and the resulting carbanion on its α -carbon acts as a nucleophile and attacks the carbonyl of the acetate bound to Cys112. The other two active site residues, His244 and Asn274, stabilize the transition state, which subsequently breaks down to yield acetoacetyl-ACP.

Second half reaction: Malonyl-ACP decarboxylation – As malonyl-ACP enters the active site, it is in an ionized state and its decarboxylation is aided by Phe205. After decarboxylation, the resulting negative charge on the thioester carbonyl of malonyl-ACP is stabilized by His244 and Asn274. A carbanion is formed on the α -carbon of malonyl-ACP that attacks the acetate bound to Cys112. The tetrahedral transition state is stabilized by an oxyanion hole formed by Cys112 and Gly306, which eventually breaks down to yield the product, 3-ketoacyl ACP.

Mutagenesis studies to determine the roles of active site residues

The roles of the conserved active site residues C112, H244 and N274 in *E. coli* KASIII were confirmed by Davies *et al.* (28) by site-directed mutagenesis of these residues to generate C112S, H244A and N274A active site mutants. The mutants showed significantly decreased condensation activity as compared to the wild-type; therefore activity was assessed for each of the transacylation and decarboxylation half-reactions to determine the role of each residue in the reaction mechanism. It was found that the C112S mutant could not carry out the transacylation; however it had a very high decarboxylation activity. In contrast, the H244A and N274A mutants could each carry out the transacylation reaction but lost their decarboxylation activities. These experiments indicate that C112 was important for transacylation while H244 and N274 played a role in decarboxylation of malonyl-ACP.

Almost at the same time as the Davies study, Smirnova and Reynolds (46) studied the KASIII from *Streptomyces glaucescens*, which makes both branched chain fatty acids (BCFA) and straight chain fatty acids (SCFA). Active site residue C122, which is important for transacylation, was mutated to C122A, C122Q and C122S. The C112Q mutant was found to have lost its transacylation ability and could only catalyze decarboxylation of malonyl-ACP, which led to the formation of acetyl-ACP. The mutants were expressed in *S. glaucescens* and

fatty acid profiles of the transformants were studied. The SCFA production in the C122Q mutant increased by 500% as compared to the wild-type, indicating an alternative pathway for the production of SCFA via an acetyl-ACP intermediate. In another study, Brown *et al.* (47) probed the role of certain key residues in the KASIII from *Mycobacterium tuberculosis* (mtFabH) by site-directed mutagenesis. Three sets of residues were mutated and effects of these mutations on the half-reactions and overall condensation reaction of mtFabH were studied. Mutations to the putative catalytic residues (i.e., C122A, H258A and N289A) decreased the condensation activity, indicating their importance in catalysis. Consistent with the earlier studies on *E. coli* KASIII, C122 was found to be important for transacylation while H258 and N289 were found to be critical for decarboxylation. The second set of mutations consisted of residues considered important in acyl-CoA binding (i.e., W42A, R46A and R161A). Interestingly, the R46A mutation indicated that R46 was more important for ACP binding rather than acyl-CoA binding. Another interesting observation reported in this study was the presence of conserved water molecules close to the active site, which suggested that these water molecules could be playing a role in deprotonation of Cys122.

Substrate diversity among KASIII enzymes

KASIII enzymes exhibit diverse substrate specificities. For example, KASIIIs from Gram-negative organisms (e.g. *E. coli*) are highly selective for acetyl-CoA and can also utilize propionyl-CoA (4), whereas KASIIIs from Gram-positive organisms, such as *Bacillus subtilis* (9), *Streptomyces glaucescens* (40), *Staphylococcus aureus* (10) and *Enterococcus faecalis* (29), show broad substrate specificities and can accept both straight chain and branched chain primers, with a higher selectivity for the latter. There are some exceptions to this apparent rule, as certain Gram-negative organisms such as *Thermus thermophilus* and *Bacteroides vulgatus* (48) produce

BCFAs, which indicates that KASIII in such organisms may prefer both straight and branched chain primers.

Earlier studies have attempted to understand whether the fatty acid profile of an organism is determined by the substrate specificity of KASIII or by the availability of appropriate initiating primers in the host organism. Choi *et al.* (9) isolated KASIII isozymes from *B. subtilis* (bsKASIIIa and bsKASIIIb), expressed these in *E. coli* and reported formation of BCFAs in these engineered *E. coli* strains. Fatty acid synthesis was reconstituted *in vitro* using a specific KASIII (*E. coli* ecKASIII or *B. subtilis* bsKASIIIa or bsKASIIIb) along with enzymes from *E. coli* that participate in the fatty acid synthesis pathway (FabD, FabG, FabI, FabZ or FabA enzymes). In the presence of isobutyryl-CoA, *B. subtilis* KASIIIa and KASIIIb catalyzed the production of BCFAs, which demonstrated that each enzyme from *E. coli* was capable of acting on branched-chain intermediates. In contrast, when KASIII from *E. coli* was added to the *in vitro* system for fatty acid synthesis, BCFA production was not detected. Taken together, these experiments suggested that KASIII from *E. coli* is the only enzyme in the *E. coli* fatty acid synthesis pathway that is incapable of utilizing branched chain primers, whereas other enzymes of the pathway could process branched chain primers. Thus, it was concluded that the fatty acid profiles observed in *B. subtilis* and *E. coli* (i.e. branched vs. straight-chain fatty acids) are determined by the substrate preference of their respective KASIII enzymes.

However, the production of BCFAs via a reconstituted *in vitro* *E. coli* fatty acid synthesis system reported by Choi and colleagues (9) were not supported by an *in vivo* study conducted by Smirnova and Reynolds (49). Smirnova and Reynolds showed that low levels of BCFAs could be produced upon feeding isobutyric acid to both an *E. coli* strain that expressed the exogenous *Streptomyces glaucescens* KASIII and an *E. coli* strain that expressed the native *E. coli* KASIII.

They further demonstrated that BCFA production could be increased in each strain by improving cellular uptake of isobutyric acid by adjustment of the acid to pH 5. Moreover, using cell-free extracts from *E. coli*, Smirnova and Reynolds demonstrated that addition of isobutyrate resulted in production of BCFA, and the further addition of the KASIII from *S. glaucescens* improved the efficiency of BCFA production four-fold. In conclusion, synthesis of SCFAs as compared to BCFAs was 10- to 15-fold higher *in-vitro* (in the *E. coli* cell-free extracts supplied with appropriate primers) and 200-fold higher *in-vivo* (in the *E. coli* strains expressing native *E. coli* KASIII and supplied with appropriate branched chain primers in the medium). Taken together, these data suggest that the KASIII from *E. coli* is capable of processing branched chain primers, however the poor transportation of isobutyrate into the *E. coli* cells results in lower BCFA production *in-vivo*. Nonetheless, this study, for the first time, established that *E. coli* KASIII is capable of processing branched chain primers, however at a lower efficiency compared to straight chain primers.

Hypotheses for the molecular basis of KASIII substrate diversity

Despite the availability of 37 crystal structures for KASIII from 12 different biological sources and functional data for numerous other KASIIIs, the underlying molecular basis of KASIII substrate specificity is not clearly understood. This is partly because only a handful of KASIIIs are both structurally and functionally characterized (*E. coli*, *S. aureus* and *M. tuberculosis* KASIIIs), which allows for the correlation of functional data with structural features. However, attempts have been made using superimposition of known KASIII crystal structures and multiple sequence alignments to predict residues and motifs affecting KASIII substrate preferences.

Structural analyses of KASIIIs from *E. coli* and *S. aureus* (17) showed that the rotamer conformation of a conserved Phe (corresponding to Phe304 of *E. coli* KASIII) differs between KASIIIs with narrow and broad substrate specificities, so it is proposed to affect the shape of the active site cleft and thereby, affecting the KASIII substrate preferences. Superposition of KASIII crystal structures revealed that this conserved Phe is oriented toward the active site (*anti* conformation) in KASIIIs that exhibit narrow substrate specificity (e.g. *E. coli* KASIII), whereas it faces away from the active site (*gauche* conformation) in KASIIIs that have broad substrate specificity (e.g. *S. aureus* KASIII) (figure 5). Two residues (corresponding to Val215 and Leu220 in *E. coli* KASIII) positioned in a layer removed from the active site vicinity are thought to orient the Phe in a particular rotamer conformation. KASIIIs that have broad substrate specificities typically have either a Phe or Trp corresponding to Val 215 in *E. coli* KASIII and Met or Ile that corresponding to the Leu220 of *E. coli* KASIII. These larger residues (Phe/Trp and Met/Ile) are thought to create steric hindrance so that the conserved Phe is forced to orient toward the active site, adopting the *gauche* conformation. This hypothesis has not been verified experimentally; its experimental validation may lead to of the discovery of the molecular basis

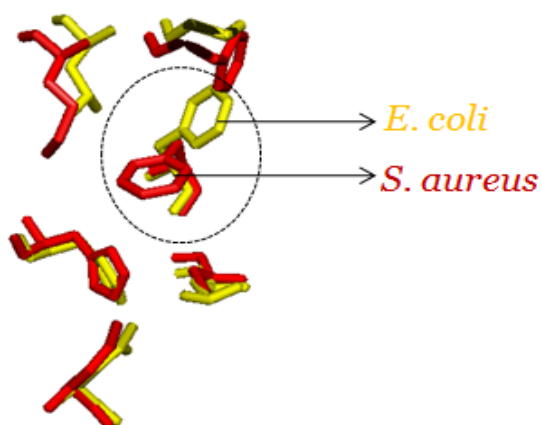


Figure 5. Different rotamer conformations of the conserved Phe residue in KASIIIs with narrow vs. broad substrate specificities. The conserved Phe adopts two different rotamer conformations – an *anti* rotamer in KASIIIs that exhibit narrow substrate specificities such as that from Gram-negative *E. coli* (F304 in yellow) and *gauche* rotamer in KASIIIs that exhibit broad substrate specificities, such as the KASIII from Gram-positive *S. aureus* (F298 in red).

for KASIII substrate specificity.

KASIII as a target for antimicrobial drugs and for synthesizing bio-based chemicals

KASIII is considered a promising target for antibacterial drug design owing to its essential role as an initiator and regulator of fatty acid biosynthesis in bacteria. Moreover, the active site residues are conserved across various bacterial species but not in mammalian systems, which presents an opportunity to block fatty acid synthesis in pathogenic bacteria while not impacting this pathway in human patients. Research on KASIII has been driven, in part, by a motivation to design novel antibacterial drugs against bacterial pathogens, including *S. aureus*, *E. coli* (50), *Pseudomonas fluorescens* (51) and *Klebsiella pneumoniae* (52). More recently, the KASIII from *Mycobacterium tuberculosis* (53) has received attention because of its potential to serve as a target for anti-tubercular drug design. Similarly, the KASIII from the malaria parasite *Plasmodium falciparum*, has become a target for the development of anti-malarial drugs (54,55), based on the recent demonstration that it possesses a type II FAS pathway.

In plants, over-expression of KASIII has been studied in order to increase the seed oil content by altering the fatty acid composition (56).

In addition to understanding the structure-function relationships of KASIII enzymes, objective of the research undertaken in this dissertation is to identify and develop novel KASIII biocatalysts that can utilize atypical substrates, such as hydroxylated, aromatic, acidic or unsaturated acyl-CoAs. Such unique KASIII will allow diversification of the fatty acid synthesis pathway to produce novel fatty acids with new functionalities at their terminal ω -ends. Such functionalized fatty acids can serve as replacements for petroleum based feedstock and can have potential applications in the current bio-based chemical industry (57).

Significance of this body of work

Scientific significance

The primary motivation dissertation research reported herein comes from the lack of understanding the structure-function relationships of KASIII and also from the realization that such an understanding can enable exquisite control over *de-novo* fatty acid synthesis. Specifically, the understanding of structure-function relationships will be immensely useful for designing novel KASIII biocatalysts with altered substrate specificities. Additionally, considering that KASIII enzymes exhibit a wide range of substrate specificities, it is important to explore the limits to this diversity by testing KASIII functionality with unique acyl-CoA substrates (e.g. cyclic-CoAs, hydroxylated-CoAs and unsaturated acyl-CoAs). Characterization of KASIIIs that act upon these unique substrates will not only enrich the existing knowledge base regarding substrate specificities but will also provide insights into structural diversity of these enzymes.

Industrial significance

This research was supported by the NSF-funded Engineering Research Center for Biorenewable Chemicals (CBiRC), and fits into the core projects of CBiRC with the ultimate aim of developing bio-based chemicals by combining biological and chemical catalysis into a united platform. The biological catalysis component is focused on the metabolic engineering of the fatty acid synthesis pathway. Because KASIII is a key regulator of this pathway and also a determinant of fatty acid type, it is an excellent candidate biocatalyst for introducing novel functionalities into the final fatty acid products (57). Novel KASIII enzymes could be used to synthesize unique fatty acids, for example fatty acids with terminal hydroxyl, amino or halogenated groups that could have applications in the bio-based polymer, surfactant and

lubricants industries. Therefore, this research on KASIII enzymes was driven in part by CBiRC's vision to find bio-based chemical replacements for petroleum-based chemicals, and also by the fundamental drive to understand the determinants of KASIII substrate specificity, so as to establish rules for rational design.

References

1. Rock, C. O., and Cronan, J. E. (1996) Escherichia coli as a model for the regulation of dissociable (type II) fatty acid biosynthesis. *Biochim Biophys Acta* **1302**, 1-16
2. Lu, Y. J., Zhang, Y. M., and Rock, C. O. (2004) Product diversity and regulation of type II fatty acid synthases. *Biochem Cell Biol* **82**, 145-155
3. Heath, R. J., and Rock, C. O. (2002) The Claisen condensation in biology. *Nat Prod Rep* **19**, 581-596
4. Tsay, J. T., Oh, W., Larson, T. J., Jackowski, S., and Rock, C. O. (1992) Isolation and characterization of the beta-ketoacyl-acyl carrier protein synthase III gene (fabH) from Escherichia coli K-12. *J Biol Chem* **267**, 6807-6814
5. Tsay, J. T., Oh, W., Larson, T. J., Jackowski, S., and Rock, C. O. (1992) Isolation and characterization of the beta-ketoacyl-acyl carrier protein synthase III gene (fabH) from Escherichia coli K-12. *J Biol Chem* **267**, 6807-6814
6. Jackowski, S., Murphy, C. M., Cronan, J. E., and Rock, C. O. (1989) Acetoacetyl-acyl carrier protein synthase. A target for the antibiotic thiolactomycin. *J Biol Chem* **264**, 7624-7629
7. Jackowski, S., and Rock, C. O. (1987) Acetoacetyl-acyl carrier protein synthase, a potential regulator of fatty acid biosynthesis in bacteria. *J Biol Chem* **262**, 7927-7931
8. Heath, R. J., and Rock, C. O. (1996) Inhibition of beta-ketoacyl-acyl carrier protein synthase III (FabH) by acyl-acyl carrier protein in Escherichia coli. *J Biol Chem* **271**, 10996-11000
9. Choi, K. H., Heath, R. J., and Rock, C. O. (2000) beta-ketoacyl-acyl carrier protein synthase III (FabH) is a determining factor in branched-chain fatty acid biosynthesis. *J Bacteriol* **182**, 365-370

10. Qiu, X., Choudhry, A. E., Janson, C. A., Grooms, M., Daines, R. A., Lonsdale, J. T., and Khandekar, S. S. (2005) Crystal structure and substrate specificity of the beta-ketoacyl-acyl carrier protein synthase III (FabH) from *Staphylococcus aureus*. *Protein Sci* **14**, 2087-2094
11. Choi, K. H., Kremer, L., Besra, G. S., and Rock, C. O. (2000) Identification and substrate specificity of beta -ketoacyl (acyl carrier protein) synthase III (mtFabH) from *Mycobacterium tuberculosis*. *J Biol Chem* **275**, 28201-28207
12. Qiu, X., Janson, C. A., Konstantinidis, A. K., Nwagwu, S., Silverman, C., Smith, W. W., Khandekar, S., Lonsdale, J., and Abdel-Meguid, S. S. (1999) Crystal structure of beta-ketoacyl-acyl carrier protein synthase III. A key condensing enzyme in bacterial fatty acid biosynthesis. *J Biol Chem* **274**, 36465-36471
13. Qiu, X., Janson, C. A., Smith, W. W., Head, M., Lonsdale, J., and Konstantinidis, A. K. (2001) Refined structures of beta-ketoacyl-acyl carrier protein synthase III. *J Mol Biol* **307**, 341-356
14. Davies, C., Heath, R. J., White, S. W., and Rock, C. O. (2000) The 1.8 Å crystal structure and active-site architecture of beta-ketoacyl-acyl carrier protein synthase III (FabH) from *Escherichia coli*. *Structure* **8**, 185-195
15. Khandekar, S. S., Konstantinidis, A. K., Silverman, C., Janson, C. A., McNulty, D. E., Nwagwu, S., Van Aller, G. S., Doyle, M. L., Kane, J. F., Qiu, X., and Lonsdale, J. (2000) Expression, purification, and crystallization of the *Escherichia coli* selenomethionyl beta-ketoacyl-acyl carrier protein synthase III. *Biochem Biophys Res Commun* **270**, 100-107
16. Daines, R. A., Pendrak, I., Sham, K., Van Aller, G. S., Konstantinidis, A. K., Lonsdale, J. T., Janson, C. A., Qiu, X., Brandt, M., Khandekar, S. S., Silverman, C., and Head, M. S. (2003) First X-ray cocrystal structure of a bacterial FabH condensing enzyme and a small molecule inhibitor achieved using rational design and homology modeling. *J Med Chem* **46**, 5-8
17. Gajiwala, K. S., Margosiak, S., Lu, J., Cortez, J., Su, Y., Nie, Z., and Appelt, K. (2009) Crystal structures of bacterial FabH suggest a molecular basis for the substrate specificity of the enzyme. *FEBS Lett* **583**, 2939-2946
18. Pereira, J. H., Goh, E. B., Keasling, J. D., Beller, H. R., and Adams, P. D. (2012) Structure of FabH and factors affecting the distribution of branched fatty acids in *Micrococcus luteus*. *Acta Crystallogr D Biol Crystallogr* **68**, 1320-1328
19. Scarsdale, J. N., Kazanina, G., He, X., Reynolds, K. A., and Wright, H. T. (2001) Crystal structure of the *Mycobacterium tuberculosis* beta-ketoacyl-acyl carrier protein synthase III. *J Biol Chem* **276**, 20516-20522

20. Musayev, F., Sachdeva, S., Scarsdale, J. N., Reynolds, K. A., and Wright, H. T. (2005) Crystal structure of a substrate complex of Mycobacterium tuberculosis beta-ketoacyl-acyl carrier protein synthase III (FabH) with lauroyl-coenzyme A. *J Mol Biol* **346**, 1313-1321
21. Sachdeva, S., Musayev, F. N., Alhamadsheh, M. M., Scarsdale, J. N., Wright, H. T., and Reynolds, K. A. (2008) Separate entrance and exit portals for ligand traffic in Mycobacterium tuberculosis FabH. *Chem Biol* **15**, 402-412
22. Choi, K. H., Heath, R. J., and Rock, C. O. (2000) beta-ketoacyl-acyl carrier protein synthase III (FabH) is a determining factor in branched-chain fatty acid biosynthesis. *J Bacteriol* **182**, 365-370
23. Khandekar, S. S., Gentry, D. R., Van Aller, G. S., Warren, P., Xiang, H., Silverman, C., Doyle, M. L., Chambers, P. A., Konstantinidis, A. K., Brandt, M., Daines, R. A., and Lonsdale, J. T. (2001) Identification, substrate specificity, and inhibition of the Streptococcus pneumoniae beta-ketoacyl-acyl carrier protein synthase III (FabH). *J Biol Chem* **276**, 30024-30030
24. Qiu, X., Choudhry, A. E., Janson, C. A., Grooms, M., Daines, R. A., Lonsdale, J. T., and Khandekar, S. S. (2005) Crystal structure and substrate specificity of the beta-ketoacyl-acyl carrier protein synthase III (FabH) from Staphylococcus aureus. *Protein Sci* **14**, 2087-2094
25. Choi, K. H., Kremer, L., Besra, G. S., and Rock, C. O. (2000) Identification and substrate specificity of beta -ketoacyl (acyl carrier protein) synthase III (mtFabH) from Mycobacterium tuberculosis. *J Biol Chem* **275**, 28201-28207
26. Rosa, D. (1973) Specificity effects in FAS in Bacillus acidocaldarius
27. Hagan Harper (1999) Flourine containing natural products
28. Davies, C., Heath, R. J., White, S. W., and Rock, C. O. (2000) The 1.8 Å crystal structure and active-site architecture of beta-ketoacyl-acyl carrier protein synthase III (FabH) from escherichia coli. *Structure* **8**, 185-195
29. Gajiwala, K. S., Margosiak, S., Lu, J., Cortez, J., Su, Y., Nie, Z., and Appelt, K. (2009) Crystal structures of bacterial FabH suggest a molecular basis for the substrate specificity of the enzyme. *FEBS Lett* **583**, 2939-2946
30. Scarsdale, J. N., Kazanina, G., He, X., Reynolds, K. A., and Wright, H. T. (2001) Crystal structure of the Mycobacterium tuberculosis beta-ketoacyl-acyl carrier protein synthase III. *J Biol Chem* **276**, 20516-20522
31. Heath, R. J., and Rock, C. O. (2002) The Claisen condensation in biology. *Nat Prod Rep* **19**, 581-596

32. Austin, M. B., and Noel, J. P. (2003) The chalcone synthase superfamily of type III polyketide synthases. *Nat Prod Rep* **20**, 79-110
33. Weng, J. K., and Noel, J. P. (2012) Structure-function analyses of plant type III polyketide synthases. *Methods Enzymol* **515**, 317-335
34. <Jackowski Rock - Acetoacetyl ACP Synthase - a regulator of FAB in bacteria_1987.pdf>.
35. Alhamadsheh, M. M., Waters, N. C., Sachdeva, S., Lee, P., and Reynolds, K. A. (2008) Synthesis and biological evaluation of novel sulfonyl-naphthalene-1,4-diols as FabH inhibitors. *Bioorg Med Chem Lett* **18**, 6402-6405
36. Lee, J. Y., Jeong, K. W., Lee, J. U., Kang, D. I., and Kim, Y. (2009) Novel E. coli beta-ketoacyl-acyl carrier protein synthase III inhibitors as targeted antibiotics. *Bioorg Med Chem* **17**, 1506-1513
37. Li, H. Q., Shi, L., Li, Q. S., Liu, P. G., Luo, Y., Zhao, J., and Zhu, H. L. (2009) Synthesis of C(7) modified chrysin derivatives designing to inhibit beta-ketoacyl-acyl carrier protein synthase III (FabH) as antibiotics. *Bioorg Med Chem* **17**, 6264-6269
38. Shi, L., Fang, R. Q., Zhu, Z. W., Yang, Y., Cheng, K., Zhong, W. Q., and Zhu, H. L. (2010) Design and synthesis of potent inhibitors of beta-ketoacyl-acyl carrier protein synthase III (FabH) as potential antibacterial agents. *Eur J Med Chem* **45**, 4358-4364
39. Huynh, K. H., Natarajan, S., Song, N. H., Ngo, P. T., Ahn, Y. J., Kim, J. G., Lee, B. M., Eo, Y. D., and Kang, L. W. (2009) Cloning, expression, crystallization and preliminary X-ray crystallographic analysis of beta-ketoacyl-ACP synthase III (FabH) from *Xanthomonas oryzae* pv. *oryzae*. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **65**, 460-462
40. Han, L., Lobo, S., and Reynolds, K. A. (1998) Characterization of beta-ketoacyl-acyl carrier protein synthase III from *Streptomyces glaucescens* and its role in initiation of fatty acid biosynthesis. *J Bacteriol* **180**, 4481-4486
41. Wen, L., Chmielowski, J. N., Bohn, K. C., Huang, J. K., Timsina, Y. N., Kodali, P., and Pathak, A. K. (2009) Functional expression of *Francisella tularensis* FabH and FabI, potential antibacterial targets. *Protein Expr Purif* **65**, 83-91
42. Khandekar, S. S., Gentry, D. R., Van Aller, G. S., Warren, P., Xiang, H., Silverman, C., Doyle, M. L., Chambers, P. A., Konstantinidis, A. K., Brandt, M., Daines, R. A., and Lonsdale, J. T. (2001) Identification, substrate specificity, and inhibition of the *Streptococcus pneumoniae* beta-ketoacyl-acyl carrier protein synthase III (FabH). *J Biol Chem* **276**, 30024-30030
43. Tai, H., Post-Beittenmiller, D., and Jaworski, J. G. (1994) Cloning of a cDNA encoding 3-ketoacyl-acyl carrier protein synthase III from *Arabidopsis*. *Plant Physiol* **106**, 801-802

44. Singh, A. K., Zhang, Y. M., Zhu, K., Subramanian, C., Li, Z., Jayaswal, R. K., Gatto, C., Rock, C. O., and Wilkinson, B. J. (2009) FabH selectivity for anteiso branched-chain fatty acid precursors in low-temperature adaptation in *Listeria monocytogenes*. *FEMS Microbiol Lett* **301**, 188-192
45. Gonzalez-Mellado, D., von Wettstein-Knowles, P., Garces, R., and Martinez-Force, E. (2010) The role of beta-ketoacyl-acyl carrier protein synthase III in the condensation steps of fatty acid biosynthesis in sunflower. *Planta* **231**, 1277-1289
46. Smirnova, N., and Reynolds, K. A. (2001) Engineered fatty acid biosynthesis in *Streptomyces* by altered catalytic function of beta-ketoacyl-acyl carrier protein synthase III. *J Bacteriol* **183**, 2335-2342
47. Brown, A. K., Sridharan, S., Kremer, L., Lindenberg, S., Dover, L. G., Sacchettini, J. C., and Besra, G. S. (2005) Probing the mechanism of the *Mycobacterium tuberculosis* beta-ketoacyl-acyl carrier protein synthase III mtFabH: factors influencing catalysis and substrate specificity. *J Biol Chem* **280**, 32539-32547
48. O'Leary W, M. (1962) The Fatty Acids of Bacteria. *Bacteriol Rev* **26**, 421-447
49. Smirnova, N., and Reynolds, K. A. (2001) Branched-chain fatty acid biosynthesis in *Escherichia coli*. *J Ind Microbiol Biotechnol* **27**, 246-251
50. Lee, J. Y., Jeong, K. W., Shin, S., Lee, J. U., and Kim, Y. (2009) Antimicrobial natural products as beta-ketoacyl-acyl carrier protein synthase III inhibitors. *Bioorg Med Chem* **17**, 5408-5413
51. Shi, L., Fang, R. Q., Zhu, Z. W., Yang, Y., Cheng, K., Zhong, W. Q., and Zhu, H. L. (2010) Design and synthesis of potent inhibitors of beta-ketoacyl-acyl carrier protein synthase III (FabH) as potential antibacterial agents. *Eur J Med Chem* **45**, 4358-4364
52. Lee, J. Y., Jeong, K. W., Lee, J. U., Kang, D. I., and Kim, Y. (2009) Novel *E. coli* beta-ketoacyl-acyl carrier protein synthase III inhibitors as targeted antibiotics. *Bioorg Med Chem* **17**, 1506-1513
53. Brown, A. K., Taylor, R. C., Bhatt, A., Futterer, K., and Besra, G. S. (2009) Platensimycin activity against mycobacterial beta-ketoacyl-ACP synthases. *PLoS One* **4**, e6306
54. Du, Y., Gisselberg, J. E., Johnson, J. D., Lee, P. J., Prigge, S. T., and Bachmann, B. O. (2010) *Lactococcus lactis* fabH, encoding beta-ketoacyl-acyl carrier protein synthase, can be functionally replaced by the *Plasmodium falciparum* congener. *Appl Environ Microbiol* **76**, 3959-3966
55. Lee, P. J., Bhonsle, J. B., Gaona, H. W., Huddler, D. P., Heady, T. N., Kreishman-Deitrick, M., Bhattacharjee, A., McCalmont, W. F., Gerena, L., Lopez-Sanchez, M., Roncal, N. E., Hudson, T. H., Johnson, J. D., Prigge, S. T., and Waters, N. C. (2009)

- Targeting the fatty acid biosynthesis enzyme, beta-ketoacyl-acyl carrier protein synthase III (PfKASIII), in the identification of novel antimalarial agents. *J Med Chem* **52**, 952-963
56. Verwoert, II, van der Linden, K. H., Walsh, M. C., Nijkamp, H. J., and Stuitje, A. R. (1995) Modification of *Brassica napus* seed oil by expression of the *Escherichia coli* *fabH* gene, encoding 3-ketoacyl-acyl carrier protein synthase III. *Plant Mol Biol* **27**, 875-886
57. Nikolau, B. J., Perera, M. A., Brachova, L., and Shanks, B. (2008) Platform biochemicals for a biorenewable chemical industry. *Plant J* **54**, 536-545

CHAPTER II

DISSECTING THE STRUCTURAL BASIS OF DIVERSE FUNCTIONALITIES OF *E. COLI* AND *B. SUBTILIS* KASIII ENZYMES USING STD NMR BASED LIGAND INTERACTIONS

A paper to be submitted to the Journal of Biological Chemistry

Shivani Garg¹, Marna Yandeu-Nelson², and Basil J. Nikolau³

¹Primary author and researcher, ²Co-advisor, ³Co-advisor and author for correspondence

Center for Biorenewable Chemicals, and the Department of Biochemistry, Biophysics and
Molecular Biology
Iowa State University, Ames, IA – 50010

Abstract

In a type II fatty acid synthase (FAS) that typically occurs in bacteria and plants, β -Ketoacyl-ACP Synthase III (KASIII) catalyzes the initial condensation reaction in this process. This enzyme exhibits a broad range of substrate specificities that specifies the chemical structure at the ω -end of the FAS fatty acid product. For example, *E. coli* KASIII acts on straight chain acyl-CoA substrates, whereas the *B. subtilis* homologs, KASIIIa and KASIIIb, can utilize both straight and branched chain acyl-CoA substrates. The structural basis for this functional diversity in KASIII enzymes remains obscure. Based on the hypothesis that the substrate specificity of these enzymes is determined by rotamer conformation of a conserved Phe residue (Phe-304 in *E. coli*; Phe-297 in *B. subtilis* KASIIIa and Phe-310 in KASIIIb), we mutated several residues that are located close to this Phe residue that could affect its rotamer conformation. The interactions of the mutated enzymes with their acyl-CoA substrates (both straight and branched chain) were studied using Saturation Transfer Difference (STD) NMR. STD NMR data indicate

that mutations of two residues that impact the rotamer conformations of this conserved Phe of *E. coli* KASIII, with the equivalent but bulkier residues found in *B. subtilis* KASIIIa promote both straight and branched acyl-CoA substrate binding by *E. coli* KASIII. Whereas, in corollary experiments in which the *B. subtilis* KASIIIa was mutated at those sites to resemble the *E. coli* KASIII enzyme, the former lost its binding affinity for both straight and branched chain acyl-CoA substrates. Similar mutations to *B. subtilis* KASIIIb were severely detrimental to its substrate binding capability. Specific catalytic activity measurements revealed partial loss of activity of these enzymes, irrespective of increase or decrease in binding of the substrate. These data collectively established the essentiality of the conserved Phe (corresponding to Phe304 of *E. coli* KASIII) and the nature of the two residues close to it for KASIII-ligand binding and catalysis. Additionally, we identified the epitopes of the straight and branched chain acyl-CoA ligands that directly interact with the KASIII enzymes. The binding epitopes were the same in both the ligands (specifically terminal adenosine group and the terminal acyl group) suggesting that the binding modes of both the ligands are similar. In conclusion, this study has enhanced our understanding of KASIII-ligand interactions and has experimentally confirmed the role of three residues critical for these interactions.

Introduction

The fatty acid biosynthesis is a diverse metabolic process that in bacteria and plants is initiated by β -Ketoacyl-ACP Synthase III (KASIII or FabH) (1), which catalyzes the initial Claisen condensation reaction of the process (2). Typically this reaction is between a starter acyl-CoA substrate and an extender substrate (e.g. malonyl-ACP) to form a β -ketoacyl-ACP

intermediate, which is first fully reduced to an acyl-ACP, and then elongated via further cycles of condensation and reduction that constitutes fatty acid biosynthesis.

KASIII has been characterized from several bacterial (3-17), protozoan (18,19) and plant species (20-26). Functionally characterized KASIII enzymes exhibit diverse substrate specificities, utilizing acyl-CoA substrates ranging from short straight chain acyl-CoAs (e.g. acetyl-CoA, propionyl-CoA (6,20,23)), branched chain acyl-CoAs (e.g. isobutyryl-CoA, anteisovaleryl-CoA (4,6,10,15,27)) to long chain acyl-CoAs (e.g. lauroyl-CoA, palmitoyl-CoA (7,28)).

By virtue of this diverse substrate specificity of KASIII, this enzyme is thought to determine some of the diversity that occurs in fatty acid profiles of different organisms, specifically the chemical structure at the ω -end of the fatty acid products (6,16,17). For example, in many Gram-positive bacteria (i.e., *Bacillus subtilis*, *Streptomyces glaucescens*, *Staphylococcus aureus*), KASIII can utilize both branched chain and straight chain acyl-CoA substrates, resulting in the production of both iso- and anteiso-branched and straight chain fatty acids (4,6,17,27). In contrast, KASIII from many Gram-negative bacteria (e.g., *Escherichia coli*) appears to prefer straight chain acyl-CoA substrates, such as acetyl-CoA, propionyl-CoA, which results in the production of straight chain even and odd carbon-numbered fatty acids (6). Based on a study showing inability of *E. coli* KASIII to accept branched chain acyl-CoAs as substrates in an *in vitro* reconstituted fatty acid synthase system, it was proposed that straight chain substrate preference of *E. coli* KASIII was the reason for absence of branched chain fatty acids in *E. coli* (6). However, a subsequent study demonstrated that in the presence of exogenously supplied branched chain substrates, *E. coli* KASIII could produce small amounts of branched chain fatty acids *in vitro*, and lack of branched chain fatty acids *in vivo* was attributed to

unavailability of suitable starting, branched chain acyl-CoA substrates (29). Another study supported the above finding by reporting *E. coli* KASIII activity with isobutyryl-CoA, albeit this activity was ten times lower than that with acetyl-CoA (27). Regardless of these differences in the reported KASIII substrate specificities, these studies demonstrate that functionality can vary greatly among KASIII enzymes.

The *E. coli* KASIII has been studied most extensively and crystallized in either the presence (5,11) or absence of the substrate acetyl-CoA (8,9,16), and also with an indole analogue inhibitor (30). Besides *E. coli* KASIII, crystal structures are also available from eight other bacteria including *Mycobacterium tuberculosis* (28,31-33), *Staphylococcus aureus* (27) and several more (13,16,17). The catalytic triads, consisting of Cys-112, His-244 and Asn-274 of the *E. coli* protein (5,8,11) and the substrate binding pockets are well conserved among KASIIIs from different organisms (10,11,16,27). Despite the availability of crystal structures for nine KASIII enzymes, the molecular basis for KASIII's diverse substrate preferences is not clearly understood.

Recently in an attempt to understand underlying structural basis for diverse KASIII functionalities, superposition of active site residues from different KASIII crystal structures revealed that a particular conserved Phe residue at the base of the active site cleft (Phe-304 in *E. coli* KASIII, and its equivalent Phe in KASIIIs from other sources) exhibits opposite rotamer conformations in KASIII from Gram-negative (*E. coli*) and Gram-positive bacteria (*Enterococcus faecalis*, *S. aureus*) (16,17). The opposite conformations of this Phe residue in Gram-negative and Gram-positive bacteria are hypothesized to confer different substrate specificities to KASIII enzymes from the two bacterial classifications (16,17). It's also been proposed that residues positioned a layer distal from the active relative to the substrate-selective

Phe residue affect the rotamer orientation of the Phe residue, namely Val-215 and Leu-220 in *E. coli* KASIII (Fig. 1A), and Phe-209 and Met-214 in *S. aureus* KASIII (16). Underlying rationale of this hypothesis is that Gram-positive bacteria, such as *S. aureus* utilize a bulky residue at the position corresponding to Val-215 in *E. coli* KASIII, forcing the conserved Phe to adopt a rare rotamer conformation that forces the phenyl side-chain into the active site (the *gauche*-rotamer), thereby affecting the active site cleft-size, and thus KASIII substrate specificity (16).

In this study, we sought to experimentally validate this postulate. Namely, we ascertained the role of these proposed more distal residues in defining the functionality of the *E. coli* KASIII, and the two KASIII homologs of *B. subtilis* (KASIIIa and KASIIIb); these latter enzymes share 45% and 38% identity with the *E. coli* KASIII, respectively (6). Specifically, we mutated the residues proposed to affect the configuration of the substrate-selective Phe residue in the *E. coli* KASIII to make it resemble *B. subtilis* KASIIIa, and vice-versa to change the KASIIIa to make it resemble the *E. coli* enzyme. The resulting mutant enzymes were characterized to determine whether the mutations affected the predicted alterations in substrate specificity in catalysis, and the ability of the enzymes to bind straight versus branched chain acyl-CoA substrates using Saturation Transfer Difference NMR (34,35).

This study provides greater understanding of the structure-function relationship of KASIII enzymes in order to structurally predict how different KASIIIs determine substrate specificity. Such predictive modeling of the KASIII enzyme should enable rational engineering of novel KASIII-based biocatalysts for production of distinct fatty acids with different ω -end chemistries, which would have a variety of applications, including new bio-based chemical molecules (36).

Materials and methods

Tertiary structure prediction of *B. subtilis* KASIII enzymes

Tertiary structures of the *B. subtilis* KASIII homologs (i.e. KASIIIa and KASIIIb) were predicted using homology modeling. An NCBI BlastP search against the PDB database was used to identify sequences that shared >40% sequence identity with KASIIIa and KASIIIb protein sequences. For KASIIIa, this was the *S. aureus* KASIII (PDB ID – 1ZOW; chain A), showing 58% sequence identity, followed by *Aquifex aeolicus* KASIII (PDB ID – 2EBD; chain A) with 52% sequence identity. We used these latter two KASIII experimentally determined structures as the template sequences for KASIIIa, and aligned each of these with the KASIIIa sequence using ClustalW alignment software (37). The target-template sequence alignments were used to model tertiary structures of KASIIIa using the alignment mode of Swiss Model (<http://swissmodel.expasy.org>) (38-40). The two different models obtained were assessed for their quality using Verify3D and Anolea (41,42), and the best model was chosen for further analysis. Similar approach was used for predicting tertiary structure of *B. subtilis* KASIIIb using as template sequences the *Thermus thermophilus* KASIII (PDB ID – 1UB7; chain A) that showed maximum sequence identity of 44%, and *S. aureus* KASIII that shared 42% identity. The PDB files of selected models were analyzed using PyMol software (www.pymol.org) (43).

Gene cloning

The *E. coli fabH* gene that encodes KASIII was PCR-amplified from *E. coli* strain MG1655 (*E. coli* Genetic Stock Center, New Haven, CT), and cloned into pDEST17 vector using Gateway cloning (Invitrogen, Carlsbad, CA), resulting in the pDEST_KASIII construct. The *yhfb* and *yjaX* genes encoding *B. subtilis* KASIIIa and KASIIIb respectively were PCR-amplified from *B. subtilis* strain 168 (*Bacillus* Genetic Stock Center, Columbus, Ohio). These

genes were cloned into pET30a expression vector (Novagen, Merck, Germany) to construct pET30_KASIIIa and pET30_KASIIIb, and they were also cloned into the pDEST17 expression vector, using Gateway cloning to generate pDEST_KASIIIa and pDEST_KASIIIb. Each pDEST17 and pET30a construct carried an N-terminal His-tagged KASIII gene sequences. The resulting plasmids were all confirmed by sequencing.

Expression and purification of recombinant wild-type and mutant KASIII proteins

E. coli OverExpressTM C41 (Lucigen, Middletown, WI) strain was used for expression of KASIII proteins from constructs pDEST_KASIII, pET30_KASIIIa, pET30_KASIIIb, pDEST_KASIIIa and pDEST_KASIIIb. The transformants were grown at 37°C in 50 ml Luria-Bertani medium, supplemented with 100 µg/ml ampicillin (Research Products International Corps., Mount Prospect, IL) for pDEST_KASIII, pDEST_KASIIIa and pDEST_KASIIIb or 50 µg/ml kanamycin (RPI Corps.) for pET30_KASIIIa and pET30_KASIIIb. The cultures were induced by addition of IPTG (Gold Biotechnology, Olivette, MO) to a final concentration of 0.4 mM when the OD₆₀₀ of the culture was between 0.6-0.8. After incubation for another 16-18 h at 25°C, cells were harvested by centrifugation (10,000 X g, 4°C, 10 min). Soluble proteins were extracted by first suspending the cell pellet in lysis buffer (0.5 M NaCl, 5 mM imidazole, 20 mM Tris-HCl, pH 8.0, 0.1 mg/ml phenylmethylsulfonyl fluoride, 0.1% Triton-X 100), followed by sonication (10 s pulses separated by 3 s intervals for a total of 3 min) and centrifugation (10,000 X g, 4°C, 30 min). The resulting supernatant (soluble protein fraction) was analyzed by SDS-PAGE. Based on the small-scale expression experiments that optimized the conditions for obtaining the highest yield of soluble recombinant KASIII proteins, we used the constructs pDEST_KASIII, pET30_KASIIIa and pDEST_KASIIIb for large-scale expression and purification of wild-type and mutant proteins. Cultures were grown, induced, harvested and

soluble protein was extracted as described in small-scale expression methods. The soluble protein fraction was filtered through a 0.45 μ filter (Corning, the Netherlands) and applied to 8 ml Ni-NTA His-bind resin. After washing the unbound protein with wash buffers I and II (0.5 M NaCl, 20 mM Tris-HCl, pH 8.0) supplemented with 20 mM and 40 mM imidazole respectively. The proteins of interest were eluted from the column with the same buffer containing 250 mM imidazole. The purified His-tagged KASIII protein preparations were dialyzed against sodium phosphate buffer, pH 7.4 and concentrated using 10,000 molecular weight cut-off ultrafiltration centrifugation filters (Millipore, Billerica, MA) at 4°C. The concentrated proteins were either supplemented with 16% glycerol and stored at -80°C or immediately used for spectrophotometric activity assays and Saturation Transfer Difference NMR experiments. Protein purity was assessed by Coomassie-staining SDS-PAGE gels, which showed presence of near-homogenous, pure proteins (greater than 95% purity). Protein concentrations were determined by Bradford's assay (BioRad, Hercules, CA).

Site-directed mutagenesis of *E. coli* and *B. subtilis* KASIII enzymes

QuikChange® Site Directed Mutagenesis Kit (Stratagene, La Jolla, CA) was used for introducing point mutations in the pDEST_KASIII, pET30_KASIIIa and pDEST_KASIIIb constructs. Four mutants were generated from each of these constructs as indicated in Table 1 using the listed primers and their complements.

Circular dichroism spectroscopy of KASIII proteins

All CD spectra of purified KASIII proteins (0.1–0.25 mg/ml in 10mM sodium phosphate buffer, pH 7.2) were collected with Jasco J-710 Spectropolarimeter, in a 0.1cm cell at 25°C. Far-UV spectra were recorded with a bandwidth of 1.0 nm and a time response of 8 s with total of 2 data accumulations.

Purification of recombinant malonyl-CoA ACP transacylase (MCAT or FabD), β -ketoacyl ACP reductase (FabG), holo-acyl carrier protein (ACP)

FabD, FabG and holo-ACP proteins are required for the spectrophotometric assay of KASIII, therefore genes encoding these proteins, (*fabD*, *fabG* and *acpP*) cloned in pCA24N expression vectors were obtained from National BioResource Project (NIG, Japan) (44). The *acpP* gene was further cloned into pETDUET vector (Novagen, Merck KGaA, Darmstadt, Germany) along with the *aas* gene (obtained from NBRP, Japan in pCA24N vector) that encodes for the fusion protein, 2-acylglycerophospho-ethanolamine acyl transferase/acyl-acyl carrier protein synthetase. These three recombinant proteins (FabD, FabG and holo-ACP), fused with N-terminal His-tag, were purified to near-homogeneity using the same procedure as described for the purification for KASIII proteins. Purity of these proteins was assessed by SDS-PAGE.

KASIII activity assays

Activity of KASIII enzymes was determined with different acyl-CoA substrates (acetyl-CoA, isobutyryl-CoA) using an enzyme-coupled spectrophotometric assay. The assay was performed in 96-well plate format with three replicates for each reaction condition. In a total volume of 100 μ l for each reaction the mix contained 100 μ M holo-ACP, 200 μ M malonyl-CoA, 10 mM DTT, 50 μ M acyl-CoA substrate (either acetyl-CoA or isobutyryl-CoA) and 200 μ M NADPH in 0.1 M sodium phosphate buffer (pH 7.2). This mix was pre-incubated with 60 μ g of MCAT (FabD) for two minutes, and the reaction was started by the addition of 30 μ g of KAR (FabG) and varying concentrations of KASIII enzyme (0.5-15 μ g). The rate of change in absorbance at 340 nm, due to the oxidation of NADPH to NADP⁺, in coupling with the reduction of 3-ketoacyl-ACP to 3-hydroxy-acyl ACP by KAR (FabG) was recorded using a Biotek multi-plate reader.

Saturation transfer difference NMR experiments

NMR experiments were performed at 25°C with a Bruker AV700 MHz spectrometer equipped with a 5 mm HCN cryoprobe. Samples for STD NMR experiments were prepared in 0.1 M sodium phosphate buffer (pH 7.2) with 5% D₂O. The protein/ligand ratio was set as 1:100 with KASIII concentration in the 20-25 µM range. Saturation was applied as CW pulse with a power of 58-60 Hz, with on-resonance pulses at 0.62 ppm and 6.86 ppm for upfield and downfield respectively. The off-resonance pulse was applied at 45 ppm, and 3-9-19 WATERGATE suppression was used with a T1-ρ filter. Saturation time of 5 s with an additional delay of 5.1 s was used. For each KASIII enzyme, a set of three replicates were acquired, with 32 scans in each replicate, for each STD NMR experiment, with acetyl-CoA or isobutyryl-CoA as ligands (Sigma-Aldrich, St. Louis, MO). The reference and the saturated spectra were obtained in an interleaved manner. Resonance assignments of ¹H NMR spectra of free ligands were completed using 1D ¹H and COSY.

Control experiments were carried out by applying on and off-resonance saturation pulses either to the ligand in the absence of protein, or to the ligand in the presence of denatured protein. Time dependence of the saturation transfer was calculated by varying the saturation times from 0.1ms to 100s, which showed that 5s was sufficient to transfer the saturation from protein to the ligand. Topspin (Bruker Biospin Corp., Billerica, MA) was used for processing the reference and saturated spectra and integrating areas of the peaks that showed STD effect.

Relative STD effects (A_{STD}) were calculated according to the equation $A_{STD} = (I_o - I_{sat})/I_o$ where I_{sat} is the signal intensity of the saturated spectrum and I_o is the signal intensity of the reference spectrum (34). STD amplification factor was calculated according to the equation, $STD_{af} = A_{STD} \times \text{molar ligand excess}$.

Competition binding experiments were performed under the conditions described above, using the first ligand at a constant concentration of 10mM, and a competing ligand was present at molar ratios of 1, 2, 4, 8 and 10-fold higher with respect to the first ligand.

Results

Differences in active site residues of *E. coli* and *B. subtilis* KASIII based on homology modeling

The tertiary structures of *B. subtilis* KASIII homologs (KASIIIa encoded by *yjaX*, and KASIIIb encoded by *yhfB*) were predicted via homology modeling using the methods described in the Experimental Procedures section. Superposition of the *E. coli* KASIII crystal structure (PDB code 3IL9) with predicted *B. subtilis* KASIIIa and KASIIIb structures allowed the identification of active site cleft residues that have different conformations in KASIII enzymes from the two organisms. A previous hypothesis suggests that the rotamer conformation of a conserved Phe (Phe-304 in *E. coli*) may impact substrate specificity of KASIII enzymes (16). In accordance with this hypothesis, we observed that while with the *E. coli* KASIII, this conserved Phe-304 residue faces away from the carbonyl group on C _{α +1} carbon (i.e., its rotamer is in the *anti*-conformation with respect to the carbonyl group) (Fig. 1A), but in the predicted structure of the *B. subtilis* KASIIIa, Phe-297 faces towards the carbonyl group on C _{α +1} carbon (i.e., it exhibits *gauche*-conformation with respect to the carbonyl group). Similarly, in the predicted structure of *B. subtilis* KASIIIb, Phe-310 rotamer is in the *gauche*-conformation. The conformation of this Phe residue correlates with the substrate specificity of the three KASIII enzymes, with the *E. coli* enzyme that shows a narrow substrate specificity having the *anti*-rotamer, whereas the two *B. subtilis* KASIIIa and KASIIIb that have a broader substrate specificity having the *gauche*-rotamer

conformation for this Phe residue. Consistent with this correlation is the observation that KASIIIs of organisms that produce branched chain fatty acids and presumably have a KASIII with a broader substrate specificity have bulky residues neighboring this substrate-selective Phe residue, and these bulky residues are assumed to force the phenyl side chain of the Phe residue to the *gauche*-rotamer conformation and orient the side chain towards the active site cleft (16,17). Homology modeling of the *B. subtilis* KASIII proteins identified these bulky residues as Phe-208 and Met-213 of KASIIIa (Fig. 2A), and Trp-221 and Val-226 of KASIIIb, corresponding to the smaller residues in *E. coli* KASIII, i.e. Val-215 and Leu-220. We postulate that these bulkier residues in the *B. subtilis* KASIII enzymes causes the conserved substrate-selective Phe *gauche*-rotamer to orient differently from the *E. coli* KASIII Phe-304's *anti*-rotamer, thereby affecting the size and hydrophobicity of the active site pockets, and subsequently imparting broader substrate specificity to the *B. subtilis* KASIII enzymes.

We reasoned therefore that if substrate specificity is governed by the orientation of the conserved Phe and its orientation is determined by the residues listed above, we hypothesize that *E. coli* KASIII, when mutated to resemble *B. subtilis* KASIII at these sites, will have broader substrate specificity for both straight and branched chain ligands (Fig. 1B). In contrast, *B. subtilis* KASIIIa and KASIIIb, when mutated to resemble *E. coli* KASIII at these sites would have narrower substrate specificity and be able to bind to only straight chain substrates (Fig. 2B). To verify these hypotheses therefore, we mutated Val-215 and Leu-220 of *E. coli* KASIII to the corresponding residues occurring in *B. subtilis* KASIIIa, i.e. Phe and Met, respectively (Fig. 1B). Two single mutants (V215F and L220M) and one double mutant (V215F_L220M) were obtained. Similarly, the homologous residues in *B. subtilis* KASIIIa were mutated to equivalent residues in *E. coli* KASIII to generate two single mutants (F208V and M213L) and a double

mutant (F208V_M213L, Fig. 2B). In *B. subtilis* KASIIIb, similar set of mutations generated two single mutants (W221V and V226L) and a double mutant (W221V_V226L).

Each of the wild-type and mutated enzymes were purified to near homogeneity with the exception of *B. subtilis* KASIIIb double mutant, which formed inclusion bodies and could not be purified. CD spectra of all purified proteins were obtained to determine if the mutants folded the same as the wild-type KASIII proteins or had a different secondary structure folding (Fig. 6).

Ligand binding epitopes for *E. coli* and *B. subtilis* KASIII enzymes

STD NMR experiments elucidated interactions between purified KASIII enzymes (both wild-type and mutated) with a straight chain substrate (acetyl-CoA) and a branched chain substrate (isobutyryl-CoA). These experiments also enabled mapping of binding epitopes on the two substrates. Relative saturation transfer to each of the binding epitopes was measured and converted to the STD amplification factor (STD_{af}), which is an indicator of ligand binding (34). Ligand epitopes with high STD_{af} values are concluded to be in close contact with the enzyme (34). STD NMR experiments with *E. coli* KASIII indicate that binding epitopes are quite similar in the acetyl-CoA and isobutyryl-CoA ligands irrespective of whether wild-type or mutant KASIII proteins were tested. With both the ligands, the H_{AP} proton, positioned at C-2 of the adenine ring of CoA, received the maximum saturation transfer (Fig. 3A), indicating that this proton is in close contact with the KASIII enzyme. This result is in agreement with previous crystallographic data (5,11), which showed the adenine ring of CoA sandwiched between Trp-32 and Arg-151 of *E. coli* KASIII. With acetyl-CoA, the second highest saturation was received by the terminal proton, H_{T} (Fig. 3A), whereas with isobutyryl-CoA, proton H_{B} , which is located on C-2 of the isobutyryl portion of the ligand and is close to the H_{T} protons, received the second highest saturation transfer (Fig. 3B). Saturation transfer to all other protons in both acetyl and

isobutyryl-CoA was very low, less than 38% relative to H_{AP} whose STD_{af} is normalized to 100%. Also, H_{AP} is on the conserved pantetheine arm in each of the acyl-CoA ligands, whereas H_T is on the variable acyl end that has been shown to interact with the active site Cys of KASIII (specifically, with Cys-112 of *E. coli* KASIII (11)).

Similar to the *E. coli* KASIII, for *B. subtilis* KASIIIa the binding epitopes of acetyl-CoA are the methyl group of the acetyl moiety, carrying H_T and the adenine moiety bearing H_{AP} . These protons received the largest saturation transfer (Fig. 3D), whereas other protons received very little saturation transfer (less than 35% relative to H_{AP}). In contrast, for isobutyryl-CoA the H_B proton received maximum saturation; H_{AP} , H_T have the second and the third highest saturation transfer respectively (Fig. 3E). Interestingly, for *B. subtilis* KASIIIb, binding epitopes were slightly different from its homolog KASIIIa and its ortholog, *E. coli* KASIII. In addition to saturation transfer to H_T and H_{AP} of acetyl-CoA, and H_T , H_B and H_{AP} of isobutyryl-CoA, a new binding epitope, H_R , was identified which received strong saturation transfer in both acetyl and isobutyryl-CoA (Fig. 3F, 3G); H_R is located on C-5 of the ribose moiety of these coenzyme A esters (Fig. 3A).

With both the *E. coli* KASIII and *B. subtilis* KASIIIa, because H_T , H_{AP} protons in acetyl-CoA, and H_T , H_B and H_{AP} in isobutyryl-CoA received strong saturation transfer, and consequently had the highest STD_{af} , we focused on differences in saturation transfer to these protons. As H_T is on the acyl end of the ligand, the amount of saturation transferred to it indicates the extent of KASIII active site's interaction with the ligand (acetyl-CoA or isobutyryl-CoA). In *B. subtilis* KASIIIb, in addition to H_T , H_{AP} in acetyl-CoA and H_T , H_B and H_{AP} in isobutyryl-CoA, we also focused on H_R for comparing the effect of mutations on substrate binding because it received strong saturation transfer in both acetyl and isobutyryl-CoAs.

Val215Phe and Leu220Met mutations of *E. coli* KASIII enhance binding with branched chain acyl-CoA substrate, but decrease catalytic activity

Wild-type *E. coli* KASIII showed binding with acetyl-CoA (Fig. 3B) with STD_{af} values of ~5.0 and ~6.0 for the H_T and H_{AP} protons respectively. In contrast, STD NMR of *E. coli* KASIII with isobutyryl-CoA as the ligand resulted in relatively lower saturation transfer, STD_{af} of 2.5 to the H_T proton and ~4.0 to the H_B proton, indicating lower binding with the branched chain ligand (Fig. 3C). The *E. coli* double mutant V215F_L220M, on the other hand, exhibited increased STD_{af} for H_T, H_B and H_{AP} as compared to the wild-type enzyme for both acetyl-CoA and isobutyryl-CoA. This suggests that the double mutant, which was mutated to resemble the broad substrate specificity enzyme *B. subtilis* KASIIIa, is showing enhanced interactions with both straight and branched chain ligands.

To identify which of the two mutations in the *E. coli* double mutant had a stronger effect on substrate binding, we examined STD amplification for each of the single mutants (i.e., *E. coli* KASIII V215F and *E. coli* KASIII L220M). As compared to the wild type enzyme the V215F mutation resulted in slight increase in STD_{af} for H_T protons of both acetyl-CoA and isobutyryl-CoA. However, the L220M mutation did not alter STD_{af} for acetyl-CoA proton H_T but significantly increased STD_{af} for H_T of isobutyryl-CoA (by ~66%) (Fig 3B, 3C). These results indicate that the substrate binding by the KASIII is primarily affected by Leu-220 and Val-215 has a lesser effect on substrate binding.

Enzymological assays of *E. coli* wild-type and mutant KASIII enzymes showed that both individual mutations at Leu-220 and Val-215 reduced catalytic activity of the enzyme (Table 2). The double mutant had a decreased specific activity with both acetyl-CoA and isobutyryl-CoA as compared to the *E. coli* wild-type KASIII, whereas the single mutants showed negligible specific

activity. In summary, these two mutations that are proposed to influence the rotamer confirmation of the conserved substrate-determining Phe residue (*anti* rotamer in *E. coli* KASIII and *gauche* rotamer in *B. subtilis* KASIIIa as depicted in Fig. 1), increased the binding of both branched chain and straight chain acyl-CoA substrates to the *E. coli* KASIII, but adversely affected the catalytic activity of the enzyme.

Phe208Val and Met213Leu mutations negatively affect substrate binding and catalytic activity in *B. subtilis* KASIIIa

STD NMR experiments with wild-type *B. subtilis* KASIIIa establish that the H_T protons of both acetyl-CoA and isobutyryl-CoA had similar STD_{af} (1.5), whereas the H_{AP} proton of isobutyryl-CoA received ~28% higher saturation as compared to the acetyl-CoA H_{AP} proton of acetyl-CoA (Fig. 3D, 3E). The double mutant F208V_M213L, which is mutated to resemble the *E. coli* KASIII at two residues neighboring the proposed substrate-determining Phe-297 (Fig. 2), showed decreased binding with acetyl-CoA and isobutyryl-CoA, as evident by significantly decreased STD_{af} for each ligand (Fig. 3D, E). Although, the single mutant *B. subtilis* KASIIIa F208V did not significantly impact binding with acetyl-CoA; for isobutyryl-CoA this mutation significantly decreased saturation transfer to the H_T and H_B protons by ~40% in comparison to the wild-type. In the case of the M213L mutant, saturation transfer decreased to both acetyl-CoA and isobutyryl-CoA, with 40% decrease in STD_{af} for H_T proton of acetyl-CoA, and 60-70% decrease in STD_{af} for H_T and H_B protons of isobutyryl-CoA.

The decreased binding of acetyl-CoA and isobutyryl-CoA ligands by the *B. subtilis* KASIIIa mutants is accompanied by decreases in catalytic activity. Individual mutations of both the residues (F208V and M213L) resulted in loss of catalytic activity with acetyl-CoA as the substrate, and decreased catalytic activity towards isobutyryl-CoA (Table 2). These data clearly

indicate that both M213L and F208V mutations decrease binding and catalytic activity of *B. subtilis* KASIIIa enzyme with both straight and branched chain substrates. However, the M213L mutation has a more pronounced negative effect on the binding of substrates by *B. subtilis* KASIIIa, especially with the branched chain substrate (isobutyryl-CoA). Phe-208 also appears to influence branched chain substrate binding, but to a lesser extent.

Trp-221 and Val-226 are critical for substrate recognition and catalytic activity of *B. subtilis* KASIIIb

The wild-type *B. subtilis* KASIIIb enzyme exhibits very high STD_{af} values for both acetyl and isobutyryl-CoA, ranging between 5.0 – 10.0, as compared to STD_{af} values of between 0.5 – 2.5 for the same ligands with *B. subtilis* KASIIIa (Fig. 3F, G). This could be either due to stronger enzyme-ligand interactions for *B. subtilis* KASIIIb or high k_{off} rate for the ligands from the active site pocket of this enzyme. The double mutant of *B. subtilis* KASIIIb (W221V_V226L), that contains mutations to resemble *E. coli* KASIII at positions 221 and 226, could not be purified due to its poor solubility. The single mutants (*B. subtilis* KASIIIb W221V and *B. subtilis* KASIIIb V226L) were purified but they did not exhibit efficient saturation transfer to either of the ligands. As indicated by the negligible STD_{af} values, the W221V mutation appears to eliminate the interactions with both acetyl-CoA and isobutyryl-CoA, (Fig. 3F, G), but the V226L mutation, resulted in less severe reduction of the STD amplification factors for both substrates.

Enzymatic assays of the wild-type and mutant *B. subtilis* KASIIIb enzymes with either acetyl-CoA or isobutyryl-CoA, indicated a very high specific activity with isobutyryl-CoA and no activity with acetyl-CoA (Table 2). The W221V mutation complete eliminated of activity

with isobutyryl-CoA, whereas V226L mutation led to approximately 50% loss of this activity with isobutyryl-CoA.

These observations, along with poor solubility of the *B. subtilis* KASIIIb double mutant, imply that both Trp-221 and Val-226 are important for substrate recognition by *B. subtilis* KASIIIb. Trp-221 residue in particular appears to be critical for binding the substrate and in catalytic functionality of *B. subtilis* KASIIIb.

The role of the conserved potential substrate-determining Phe of KASIIIs

The role of the conserved Phe residue that is thought to determine substrate specificity of KASIII enzymes was investigated by directly mutating this residue to Ala. The *B. subtilis* KASIIIb F310A mutant could not be purified owing to poor expression and solubility, but mutant proteins were recovered and characterized with the *E. coli* KASIII and *B. subtilis* KASIIIa. With the *E. coli* enzyme the F304A mutation lowers the saturation transfer to the H_T protons of acetyl-CoA and isobutyryl-CoA by 60% (Fig. 4A, B), and with the *B. subtilis* enzyme the equivalent mutation (F297A) increases the STD amplification factors of protons H_T and H_{AP} for both acetyl-CoA and isobutyryl-CoA. These mutations eliminate all enzymatic activity of the *E. coli* KASIII (Table 2), but in the case of the *B. subtilis* KASIIIa, about 25% of the catalytic activity with the isobutyryl-CoA substrate is retained, but all activity with acetyl-CoA is eliminated (Table 2).

Competition binding experiments reveal relative affinities of ligands by KASIII enzymes

Competition binding experiments to determine the relative affinities of the two substrate ligands (acetyl-CoA and isobutyryl-CoA) to the KASIII enzymes. The concentration of acetyl-CoA was kept constant and isobutyryl-CoA was titrated at increasing concentrations, and saturation transfer to the both the ligands was assayed to assess if the competing ligand replaced

the acetyl-CoA ligand. As evidenced by the fact that 4 times higher concentration of isobutyryl-CoA is needed to displace acetyl-CoA, this experiment indicates that acetyl-CoA is the preferred substrate for *E. coli* KASIII (Fig. 5A).

Similar experiments for *B. subtilis* KASIIIa and *B. subtilis* KASIIIb were conducted, which indicate that isobutyryl-CoA is the preferred substrate for *B. subtilis* KASIIIa and KASIIIb; with both enzymes, a higher STD amplification factor was observed for the H_T proton of isobutyryl-CoA than that of acetyl-CoA H_T (Fig. 5B and 5C).

Discussion

Key residues that govern substrate binding and catalytic activity of *E. coli* and *B. subtilis* KASIII enzymes

Based on crystal structures of several KASIII enzymes from Gram-positive and Gram-negative bacteria in previous studies, several residues were postulated to be influencing the rotamer conformation of a conserved Phe (which is usually *anti* with respect to the carbonyl group on C_{α+1} carbon in Gram-negative bacteria's KASIII, and *gauche* in Gram-positive bacteria's KASIII), thereby altering the substrate specificities of KASIII enzymes (16,17). In this study, effects of mutations of those residues on KASIII substrate binding and catalytic activity were studied using STD NMR. Our STD NMR data indicated that in *E. coli* KASIII, residues Val-215 and Leu-220, when mutated to corresponding residues in *B. subtilis* KASIIIa i.e. Phe and Met respectively, result in increased branched chain substrate binding. By posing steric hindrance, the new bulky substitutes i.e. Phe and Met could be switching the rotamer conformation of the conserved Phe-304 of *E. coli* KASIII from *anti* to *gauche*, thereby promoting branched chain substrate binding. However, increased binding did not result in

increased catalytic activity with isobutyryl-CoA as evidenced by the specific activity data (Table 2). This implies that although mutations to *E. coli* KASIII resulted in higher binding with the substrates, but the substrate pocket environment was altered resulting in either poor catalysis of the bound substrate or failure to release the final product.

In case of *B. subtilis* KASIIIa, Met-213 appeared to have stronger negative effect on both branched and straight chain substrate binding as compared to Phe-208 which seemed to affect mainly branched chain binding. Specific activity data indicated that both the mutations result in loss of activity with straight chain substrate and decrease activity with the branched chain substrate. Binding and activity data in conjunction implies that while both the residues, Met-213 and Phe-208, are important for catalysis, they have different roles in substrate binding. Specifically, while Phe-208 is required for branched chain substrate binding, the Met-213 residue is essential for both branched and straight chain substrate binding.

Insolubility of *B. subtilis* KASIIIb double mutant, poor saturation transfer and poor catalytic activity displayed by its single mutants suggest that both Trp-221 and Val-226 are indispensable for the enzyme's functionality, and of the two residues, Trp-221 seems to influence substrate binding and catalytic activity more.

To determine if the conserved Phe is critical for substrate recognition, we mutated it to Ala in all the three wild-type KASIII enzymes. In *E. coli* KASIII, F304A mutant resulted in decreased saturation transfer to both acetyl and isobutyryl-CoA substrates; and also resulted in complete loss of catalytic activity, indicating that Phe-304 is required for substrate binding and catalysis. Whereas in *B. subtilis* KASIIIa, F297A mutant enzyme resulted in increased saturation transfer for both straight and branched chain ligands. Increase in binding was not accompanied by increase in catalysis implying that F297 is required for catalytic activity of *B. subtilis*

KASIIIa. Poor expression and solubility of *B. subtilis* KASIIIb F310A mutant, which hindered its successful purification, affirms the importance of Phe-310 in maintaining this enzyme's functionality.

Binding mode of straight chain and branched chain ligands appears to be similar

Binding epitopes of acetyl-CoA (the straight chain ligand) and isobutyryl-CoA (the branched chain ligand) were found to be the adenine moiety on the conserved CoA part of the ligands, and the terminal methyl groups on the variable acyl ends. Irrespective of the enzyme tested, adenine moiety received maximum saturation transfer suggesting that it has the closest contact with KASIII enzymes. Crystallographic data from previous studies, showing that ADP ring of acetyl-CoA is sandwiched between two conserved KASIII residues (Trp-32 and Arg-151 of *E. coli* KASIII) and forms hydrogen bonding with carbonyls of Thr-28 and Arg-151 in *E. coli* KASIII, supports our data (5,11). From our data, similarity observed in the binding epitopes implies a similar binding mechanism of branched chain ligand and straight chain ligands to the enzyme; wherein coenzyme A tail tethers to the protein via interlocking of adenine moiety with the conserved KASIII residues (Trp-32 and Arg-151 in *E. coli* KASIII (5,11)) leaving rest of the pantothenate chain swinging loosely in the CoA tunnel of the enzyme. Lack of saturation transfer to the protons of pantothenate chain corroborates this hypothesis, and is consistent with previous findings that the pantothenate chain does not make strong contacts with the hydrophobic tunnel that leads to the KASIII active site (2). Strong saturation transfer to the terminal methyl groups of the ligands is expected as these methyl groups are close to the catalytic triad of KASIII enzymes for enabling condensation of acyl group of the ligands with malonyl-ACP.

A similar binding mode of both the acyl-CoA ligands in which the CoA ring is stacked between Trp and Arg of the KASIII enzymes suggests that first the CoA ring binds to the

enzyme and possibly induces a conformational change that allows the rest of the pantothenate arm to enter the active site tunnel.

Loss of catalytic activity by the mutants of *E. coli* and *B. subtilis* KASIII enzymes suggests that the conserved Phe, and the residues neighboring to it either aid in catalysis or aid in binding of the second substrate, malonyl-ACP.

***E. coli* KASIII shows binding and activity with branched chain acyl-CoA**

Previous studies (6,27,29) have attempted to determine whether the fatty acid profile of an organism is determined by the substrate specificity of KASIII or by the availability of appropriate priming substrates (straight or branched chain acyl-CoAs) *in vivo*. As per conclusions of one study (6), *E. coli* KASIII is thought to be incapable of utilizing branched chain acyl-CoAs thereby resulting in only straight chain fatty acid production. In contrast, other studies (27,29) show that *E. coli* KASIII is capable of using branched chain primers, though at a lower efficiency compared to straight chain primers, suggesting that absence of branched primers *in vivo* hinders branched chain fatty acid production. Our STD NMR and activity assay data provides strong evidence that *E. coli* KASIII is capable of binding and processing isobutyryl-CoA, however the binding affinity is very low as compared to that for acetyl-CoA as evident in competition binding experiments. Additionally, catalytic activity for isobutyryl-CoA is approximately 10 times low in comparison to acetyl-CoA, which is comparable to the report demonstrating branched chain fatty acids production is 10-15 folds lower than straight chain fatty acid production in cell-free extracts of *E. coli* to which appropriate straight and branched chain primers were added (29). An explanation for no production of branched chain fatty acids *in vivo* by *E. coli* KASIII in a previous study (6) could be low concentration of isobutyryl-CoA than what would be physiologically required by *E. coli* KASIII *in vivo*. Therefore, based on our

results that show that *E. coli* KASIII can bind (Fig. 3) and catalyze isobutyryl-CoA (Table 2), the lack of branched chain fatty acid production in *E. coli* may be explained by absence or low availability of isobutyryl-CoA primers *in vivo*.

Functional and structural divergence of *B. subtilis* KASIII homologs

B. subtilis has two KASIII homologs and the significance of this KASIII gene redundancy has not been identified yet. In our experiments, binding epitopes of *B. subtilis* KASIIIb were found to be slightly different from KASIIIa, which suggests that *B. subtilis* KASIIIa folds quite differently from KASIIIb and is more divergent from its *E. coli* counterpart. KASIIIb has much lower sequence similarity with *E. coli* KASIII (48%, (6)) as compared to *B. subtilis* KASIIIa, which shares 54% similarity with *E. coli* KASIII (6). Furthermore, the residues that appeared to be critical for substrate recognition in KASIIIb were not exactly equivalent to the residues in KASIIIa. In KASIIIa, Met-213 seems to be more important for substrate binding whereas in KASIIIb, Trp-221 appears to have major role in substrate binding and this residue corresponds to Phe-208 of *B. subtilis* KASIIIa. All these observations strongly imply that the two *B. subtilis* homologs have diverged evolutionarily to have distinct structures and different functionalities that appear to be non-redundant. Inability to purify the F310A and double mutant of *B. subtilis* KASIIIb also indicates that this enzyme is quite different from KASIIIa. Structural analyses of these homologs would provide interesting insights into the structural differences that exist.

To conclude, this study has provided new insights into binding interactions of straight and branched chain acyl-CoA substrates with KASIII enzymes, and identified and experimentally confirmed the role of key residues that determine the substrate specificity in *E. coli* and *B. subtilis* KASIII homologs. A deeper understanding of KASIII structure-function

relationships can enable rational engineering of novel KASIII biocatalysts with unique substrate specificities. These biocatalysts can have potential applications in production of value-added bio-based chemicals.

Footnotes

The abbreviations used are: KASIII: β -Ketoacyl ACP Synthase III; STD: Saturation Transfer Difference; Phe: Phenylalanine; Val: Valine; Leu: Leucine; Met: Methionine; Trp: Tryptophan

Acknowledgements

This work was supported by NSF grant and Center for Biorenewable Chemicals. We would like to acknowledge CBiRC Research Experience for Undergraduates (REU) program students, Kim Xuan and Luke Prest, for their help with protein purification and site directed mutagenesis, and Morgan Becker for her help with the CD experiments.

Table 1. Primers for site-directed mutagenesis of *E. coli* and *B. subtilis* KASIII enzymes

Gene	Construct	Mutation	QuikChange Forward Primer (5'-3')	QuikChange Reverse Primer (5'-3')
<i>E. coli fabH</i> (<i>KASIII</i>)	pDEST_K ASIII	V215F	GCAACGAAGTCTTCAAGTTTGCGGT AACGGAAGT	CAGTTCCGTTACCGCAAACCTTGA AGACTTCGTTGC
		L220M	TTCAAGGTTGCGGTAACGGAAATG GCGCACATC	GATGTGCGCCATTTCGTTACCG CAACCTTGAA
		V215F_L 220M	GCAACGAAGTCTTCAAGTTTGCGGT AACGGAAGT TTCAAGGTTGCGGTAACGGAAATG GCGCACATC	CAGTTCCGTTACCGCAAACCTTGA AGACTTCGTTGC GATGTGCGCCATTTCGTTACCG CAACCTTGAA
		F304A	GGTTCTGCTTGAAGCCGCTGGCGGT GGATTCACC	GGTGAATCCACCGCCAGCGGCTT CAAGCAGAACC
<i>B. subtilis yjaX</i> (<i>KASIIIa</i>)	pET30_K ASIIIa	F208V	GAATGGACGAGAAGTTTTCAAAGT TGCAGTCCGCC	GGCGGACTGCAACTTTGAAAACCT TCTCGTCCATTC
		M213L	CAAATTTGCAGTCCGCCAATTGGGA GAATCATGCG	CGCATGATTCTCCCAATTGGCGG ACTGCAAATTTG
		F208V_M 213L	GAATGGACGAGAAGTTTTCAAAGT TGCAGTCCGCC CAAAGTTGCAGTCCGCCAATTGGG AGAATCATGCG	CGCATGATTCTCCCAATTGGCGG ACTGCAAATTTG CGCATGATTCTCCCAATTGGCGG ACTGCAACTTTG
		F297A	GGTCGTTATGGTAGGGGCCGCGG AGGACTAACA	TGTTAGTCCTCCGCCGGCCCCTA CCATAACGACC
<i>B. subtilis yhfB</i> (<i>KASIIIb</i>)	pDEST_K ASIIIb	W221V	GCAAAACGGACGCGAGGTATATAA AGTGGCCGCAAGAACC	GGTTCTTGCGGCCACTTTATATA CCTCGCGTCCGTTTTGC
		V226L	GGCCGCAAGAACCCTCCCTGGCGA ATT	AATTCGCCAGGGAGGGTTCTTGC GGCC
		W221VV 226L	GCAAAACGGACGCGAGGTATATAA AGTGGCCGCAAGAACC GGCCGCAAGAACCCTCCCTGGCGA ATT	GGTTCTTGCGGCCACTTTATATA CCTCGCGTCCGTTTTGC AATTCGCCAGGGAGGGTTCTTGC GGCC
		F310A	AATCGTTTTGCTTTTCGGGGCTGGC GGCGGATTAACCTAT	ATAGGTTAATCCGCCGCCAGCCC CGAAAAGCAAAACGATT

Table 2. Specific activity (nmol/min/mg) of wild-type and mutant KASIII with straight chain (acetyl) and branched chain (isobutyryl) acyl-CoA substrates.

Enzyme	Mutation	Acetyl-CoA	Isobutyryl-CoA
<i>E. coli</i> KASIII	Wild-type	447 ± 68	36 ± 14
	V215F	Not detectable	Not detectable
	L220M	Not detectable	Not detectable
	V215F_L220M	221 ± 104	18 ± 5
	F304A	Not detectable	Not detectable
<i>B. subtilis</i> KASIIIa	Wild-type	64 ± 32	205 ± 139
	F208V	Not detectable	Not detectable
	M213L	Not detectable	156 ± 47
	F208V_M213L	Not detectable	17 ± 7
	F297A	Not detectable	56 ± 28
<i>B. subtilis</i> KASIIIb	Wild-type	Not detectable	279 ± 8
	W221V	Not detectable	Not detectable
	V226L	Not detectable	134 ± 6

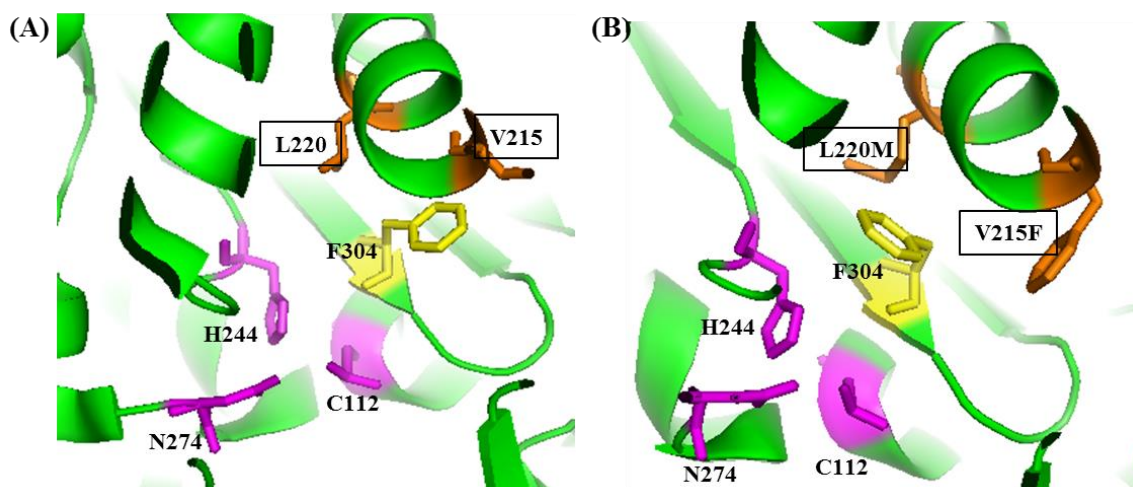


Figure 1. Proposed effect of mutations on Phe-304 orientation and on substrate specificity in *E. coli* KASIII. *E. coli* KASIII active site consists of C112, N274 and H244 (highlighted in purple) (PDB 3IL9 in (A)). F304 (highlighted in yellow) is close to the active site and is proposed to be affecting *E. coli* KASIII's substrate specificity based on its orientation owing to V215 and L220 in the layer behind it. We propose that the mutations L220M and V215F (mutated residues modeled using PyMol (B)) will affect the orientation of F304 resulting in broad substrate specificity of *E. coli* KASIII enzyme.

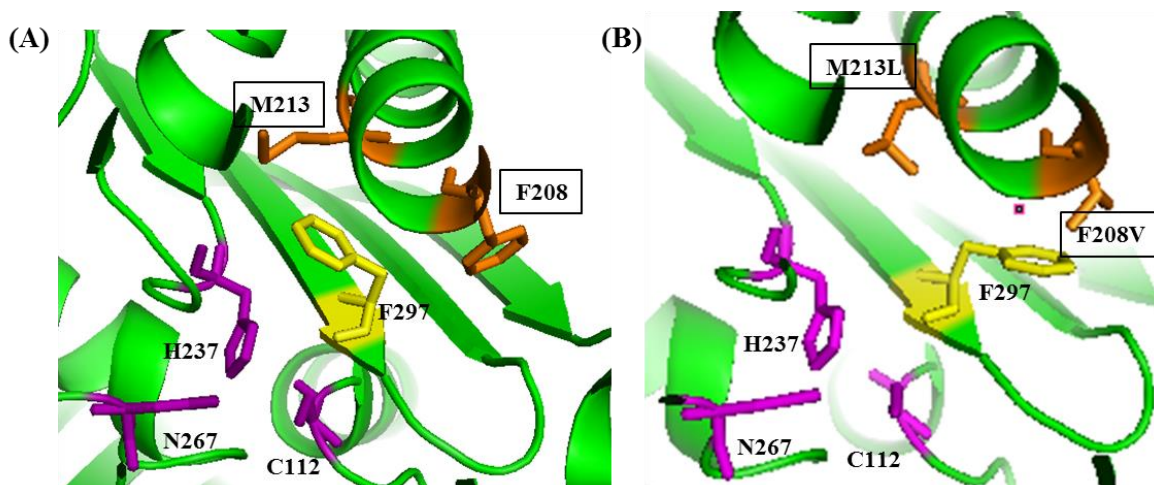


Figure 2. Proposed effect of mutations on Phe-297 orientation and on substrate specificity in *B. subtilis* KASIIIa's modeled structure. Residues C122, N267 and H237 (predicted structure of *B. subtilis* KASIIIa using Swiss Model in (A)) are the active site residues. F297 has an orientation opposite to that of *E. coli* KASIII's F304. F297 is proposed to be affecting the enzyme's substrate specificity based on its orientation owing to M213 and F208 in the layer behind it. We hypothesize that the mutations M213L and F208V (mutated residues modeled using PyMol (B)) will affect the orientation of F297 resulting in narrow substrate specificity of *B. subtilis*

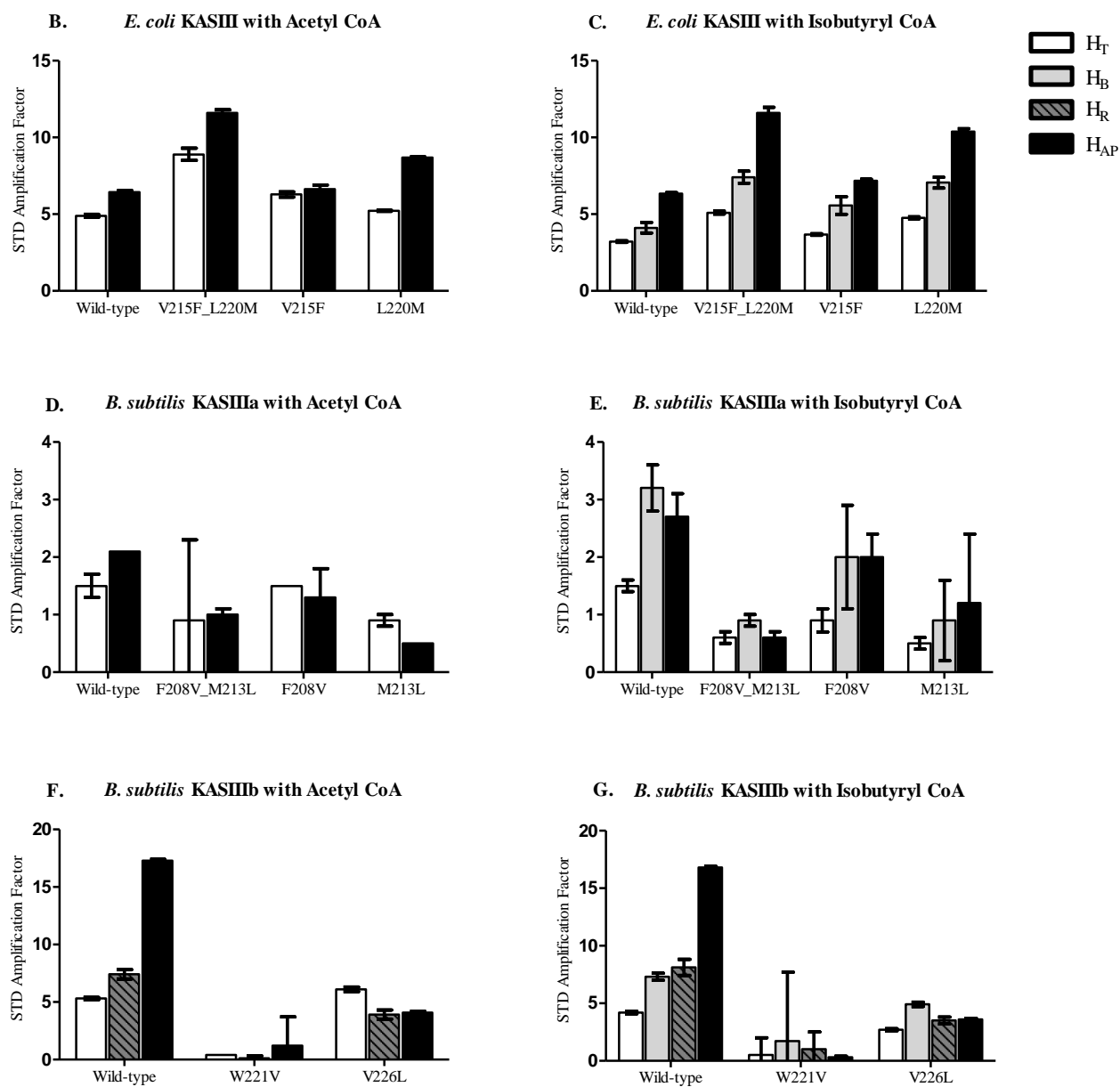
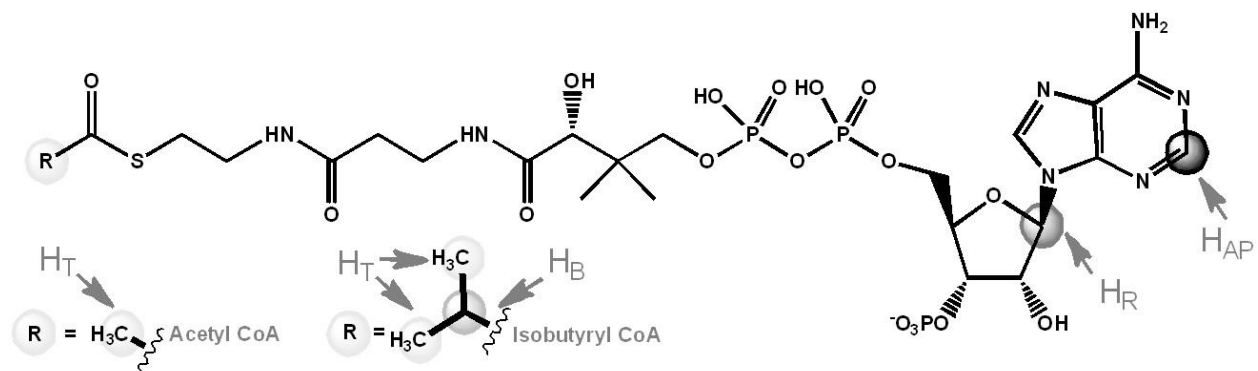


Figure 3. Effect of mutations on KASIII interaction with straight and branched chain acyl-CoA ligands as elucidated by STD NMR data. Protons on the main binding epitopes for acetyl and isobutyryl-CoA with KASIII wild-types and mutants are highlighted (A). STD amplification factors for protons receiving the maximum saturation transfer is plotted for different alleles of *E. coli* KASIII with acetyl-CoA (B) and isobutyryl-CoA (C); for different alleles of *B. subtilis* KASIIIa with acetyl-CoA (D) and isobutyryl-CoA (E); and for different alleles of *B. subtilis* KASIIIb with acetyl-CoA (E) and isobutyryl-CoA (F). Data points are average of three experiments. Each data point is average of three independent experiments.

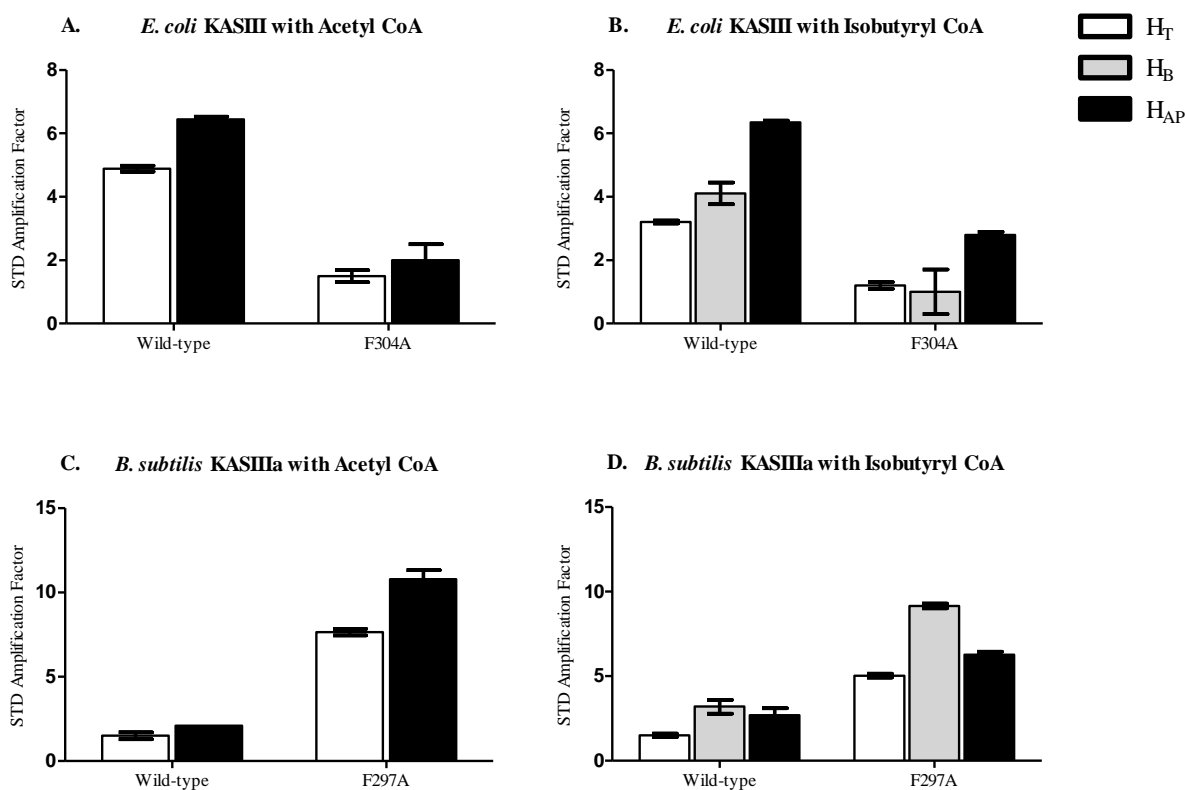


Figure 4. Effect of mutating the conserved Phe to Ala on substrate binding. For *E. coli* KASIII wild-type and F304A mutant, STD amplification factors for H_T and H_{AP} protons of acetyl and isobutyryl-CoA are shown in (A) and (B). For *B. subtilis* KASIIIa wild-type and F297A mutant, STD amplification factors for H_T and H_{AP} protons of acetyl and isobutyryl-CoA are shown in (C) and (D).

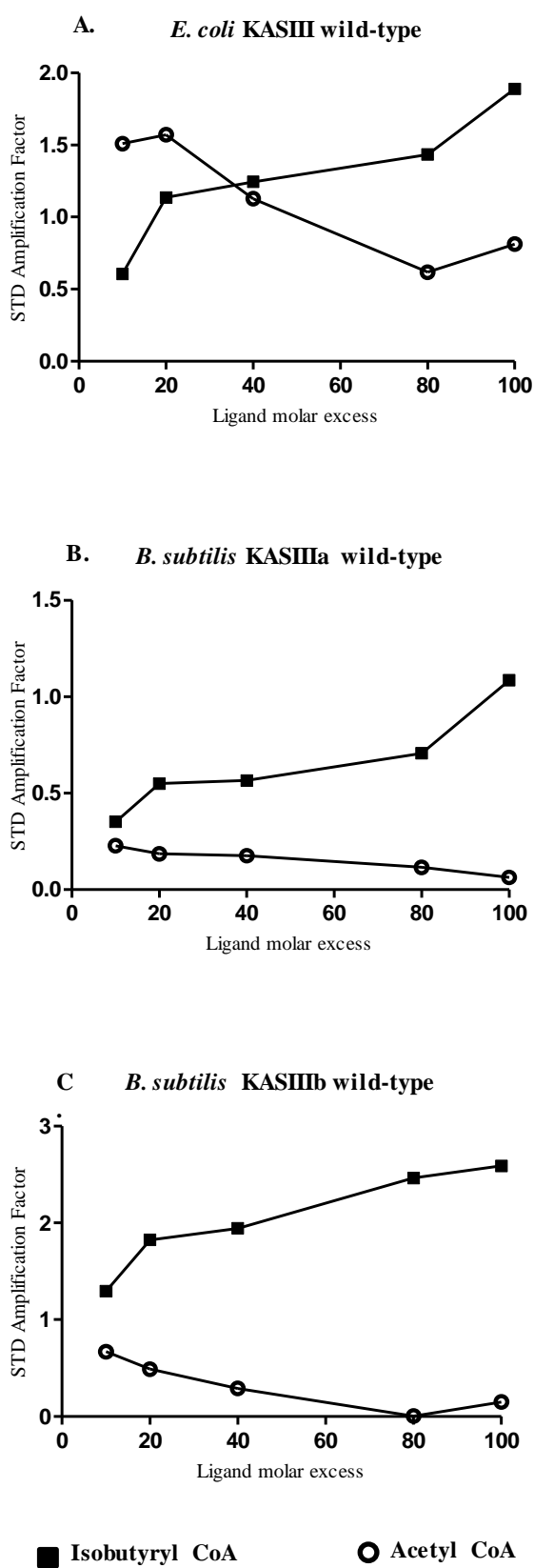


Figure 5. Competition binding data for *E. coli* KASIII, *B. subtilis* KASIIIa and KASIIIb. Displacement of one ligand by another in the binding pocket of KASIII enzyme was observed by competition binding experiments for *E. coli* KASIII wild-type (A), *B. subtilis* KASIIIa wild-type (B) and *B. subtilis* KASIIIb wild-type (C).

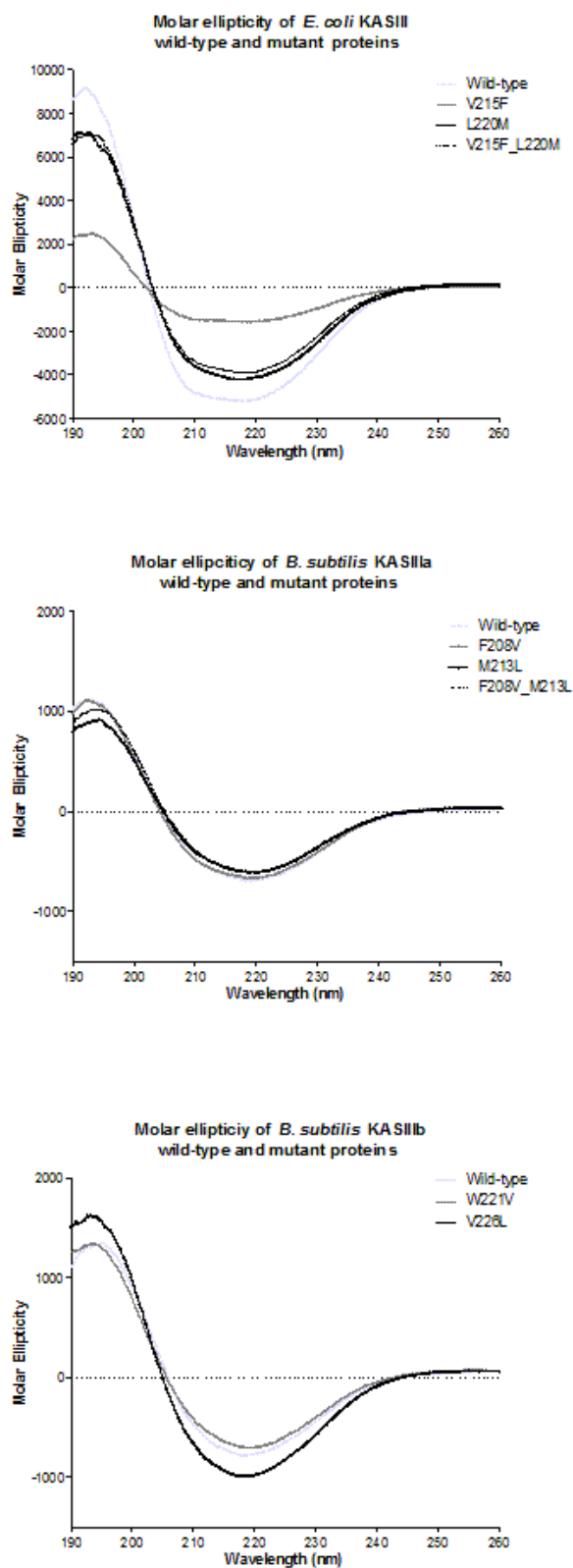


Figure 6. CD spectra of *E. coli* KASIII, *B. subtilis* KASIIIa, KASIIIb wild-type and mutant proteins. To confirm if the mutated proteins were folding the same as wild-type, Circular Dichroism spectra were obtained for *E. coli* KASIII wild-type and mutants (A), *B. subtilis* KASIIIa wild-type and mutants (B), and *B. subtilis* KASIIIb wild-type and mutants (C).

References

1. Lu, Y. J., Zhang, Y. M., and Rock, C. O. (2004) Product diversity and regulation of type II fatty acid synthases. *Biochem Cell Biol* **82**, 145-155
2. Heath, R. J., and Rock, C. O. (2002) The Claisen condensation in biology. *Nat Prod Rep* **19**, 581-596
3. Tsay, J. T., Oh, W., Larson, T. J., Jackowski, S., and Rock, C. O. (1992) Isolation and characterization of the beta-ketoacyl-acyl carrier protein synthase III gene (fabH) from *Escherichia coli* K-12. *J Biol Chem* **267**, 6807-6814
4. Han, L., Lobo, S., and Reynolds, K. A. (1998) Characterization of beta-ketoacyl-acyl carrier protein synthase III from *Streptomyces glaucescens* and its role in initiation of fatty acid biosynthesis. *J Bacteriol* **180**, 4481-4486
5. Qiu, X., Janson, C. A., Konstantinidis, A. K., Nwagwu, S., Silverman, C., Smith, W. W., Khandekar, S., Lonsdale, J., and Abdel-Meguid, S. S. (1999) Crystal structure of beta-ketoacyl-acyl carrier protein synthase III. A key condensing enzyme in bacterial fatty acid biosynthesis. *J Biol Chem* **274**, 36465-36471
6. Choi, K. H., Heath, R. J., and Rock, C. O. (2000) beta-ketoacyl-acyl carrier protein synthase III (FabH) is a determining factor in branched-chain fatty acid biosynthesis. *J Bacteriol* **182**, 365-370
7. Choi, K. H., Kremer, L., Besra, G. S., and Rock, C. O. (2000) Identification and substrate specificity of beta -ketoacyl (acyl carrier protein) synthase III (mtFabH) from *Mycobacterium tuberculosis*. *J Biol Chem* **275**, 28201-28207
8. Davies, C., Heath, R. J., White, S. W., and Rock, C. O. (2000) The 1.8 Å crystal structure and active-site architecture of beta-ketoacyl-acyl carrier protein synthase III (FabH) from *Escherichia coli*. *Structure* **8**, 185-195
9. Khandekar, S. S., Konstantinidis, A. K., Silverman, C., Janson, C. A., McNulty, D. E., Nwagwu, S., Van Aller, G. S., Doyle, M. L., Kane, J. F., Qiu, X., and Lonsdale, J. (2000) Expression, purification, and crystallization of the *Escherichia coli* selenomethionyl beta-ketoacyl-acyl carrier protein synthase III. *Biochem Biophys Res Commun* **270**, 100-107
10. Khandekar, S. S., Gentry, D. R., Van Aller, G. S., Warren, P., Xiang, H., Silverman, C., Doyle, M. L., Chambers, P. A., Konstantinidis, A. K., Brandt, M., Daines, R. A., and Lonsdale, J. T. (2001) Identification, substrate specificity, and inhibition of the *Streptococcus pneumoniae* beta-ketoacyl-acyl carrier protein synthase III (FabH). *J Biol Chem* **276**, 30024-30030
11. Qiu, X., Janson, C. A., Smith, W. W., Head, M., Lonsdale, J., and Konstantinidis, A. K. (2001) Refined structures of beta-ketoacyl-acyl carrier protein synthase III. *J Mol Biol* **307**, 341-356

12. Revill, W. P., Bibb, M. J., Scheu, A. K., Kieser, H. J., and Hopwood, D. A. (2001) Beta-ketoacyl acyl carrier protein synthase III (FabH) is essential for fatty acid biosynthesis in *Streptomyces coelicolor* A3(2). *J Bacteriol* **183**, 3526-3530
13. Huynh, K. H., Natarajan, S., Song, N. H., Ngo, P. T., Ahn, Y. J., Kim, J. G., Lee, B. M., Eo, Y. D., and Kang, L. W. (2009) Cloning, expression, crystallization and preliminary X-ray crystallographic analysis of beta-ketoacyl-ACP synthase III (FabH) from *Xanthomonas oryzae* pv. *oryzae*. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **65**, 460-462
14. Wen, L., Chmielowski, J. N., Bohn, K. C., Huang, J. K., Timsina, Y. N., Kodali, P., and Pathak, A. K. (2009) Functional expression of *Francisella tularensis* FabH and FabI, potential antibacterial targets. *Protein Expr Purif* **65**, 83-91
15. Singh, A. K., Zhang, Y. M., Zhu, K., Subramanian, C., Li, Z., Jayaswal, R. K., Gatto, C., Rock, C. O., and Wilkinson, B. J. (2009) FabH selectivity for anteiso branched-chain fatty acid precursors in low-temperature adaptation in *Listeria monocytogenes*. *FEMS Microbiol Lett* **301**, 188-192
16. Gajiwala, K. S., Margosiak, S., Lu, J., Cortez, J., Su, Y., Nie, Z., and Appelt, K. (2009) Crystal structures of bacterial FabH suggest a molecular basis for the substrate specificity of the enzyme. *FEBS Lett* **583**, 2939-2946
17. Pereira, J. H., Goh, E. B., Keasling, J. D., Beller, H. R., and Adams, P. D. (2012) Structure of FabH and factors affecting the distribution of branched fatty acids in *Micrococcus luteus*. *Acta Crystallogr D Biol Crystallogr* **68**, 1320-1328
18. Waters, N. C., Kopydlowski, K. M., Guszczynski, T., Wei, L., Sellers, P., Ferlan, J. T., Lee, P. J., Li, Z., Woodard, C. L., Shallom, S., Gardner, M. J., and Prigge, S. T. (2002) Functional characterization of the acyl carrier protein (PfACP) and beta-ketoacyl ACP synthase III (PfKASIII) from *Plasmodium falciparum*. *Mol Biochem Parasitol* **123**, 85-94
19. Prigge, S. T., He, X., Gerena, L., Waters, N. C., and Reynolds, K. A. (2003) The initiating steps of a type II fatty acid synthase in *Plasmodium falciparum* are catalyzed by pfACP, pfMCAT, and pfKASIII. *Biochemistry* **42**, 1160-1169
20. Clough, R. C., Matthis, A. L., Barnum, S. R., and Jaworski, J. G. (1992) Purification and characterization of 3-ketoacyl-acyl carrier protein synthase III from spinach. A condensing enzyme utilizing acetyl-coenzyme A to initiate fatty acid synthesis. *J Biol Chem* **267**, 20992-20998
21. Jaworski, J. G., Clough, R. C., and Barnum, S. R. (1989) A Cerulenin Insensitive Short Chain 3-Ketoacyl-Acyl Carrier Protein Synthase in *Spinacia oleracea* Leaves. *Plant Physiol* **90**, 41-44

22. Tai, H., Post-Beittenmiller, D., and Jaworski, J. G. (1994) Cloning of a cDNA encoding 3-ketoacyl-acyl carrier protein synthase III from Arabidopsis. *Plant Physiol* **106**, 801-802
23. Abbadi, A., Brummel, M., Schütt, B. S., Slabaugh, M. B., Schuch, R., and Spener, F. (2000) Reaction mechanism of recombinant 3-oxoacyl-(acyl-carrier-protein) synthase III from *Cuphea wrightii* embryo, a fatty acid synthase type II condensing enzyme. *Biochem J* **345 Pt 1**, 153-160
24. Dehesh, K., Tai, H., Edwards, P., Byrne, J., and Jaworski, J. G. (2001) Overexpression of 3-ketoacyl-acyl-carrier protein synthase IIIs in plants reduces the rate of lipid synthesis. *Plant Physiol* **125**, 1103-1114
25. Li, J., Li, M. R., Wu, P. Z., Tian, C. E., Jiang, H. W., and Wu, G. J. (2008) Molecular cloning and expression analysis of a gene encoding a putative beta-ketoacyl-acyl carrier protein (ACP) synthase III (KAS III) from *Jatropha curcas*. *Tree Physiol* **28**, 921-927
26. González-Mellado, D., von Wettstein-Knowles, P., Garcés, R., and Martínez-Force, E. (2010) The role of beta-ketoacyl-acyl carrier protein synthase III in the condensation steps of fatty acid biosynthesis in sunflower. *Planta* **231**, 1277-1289
27. Qiu, X., Choudhry, A. E., Janson, C. A., Grooms, M., Daines, R. A., Lonsdale, J. T., and Khandekar, S. S. (2005) Crystal structure and substrate specificity of the beta-ketoacyl-acyl carrier protein synthase III (FabH) from *Staphylococcus aureus*. *Protein Sci* **14**, 2087-2094
28. Scarsdale, J. N., Kazanina, G., He, X., Reynolds, K. A., and Wright, H. T. (2001) Crystal structure of the *Mycobacterium tuberculosis* beta-ketoacyl-acyl carrier protein synthase III. *J Biol Chem* **276**, 20516-20522
29. Smirnova, N., and Reynolds, K. A. (2001) Branched-chain fatty acid biosynthesis in *Escherichia coli*. *J Ind Microbiol Biotechnol* **27**, 246-251
30. Daines, R. A., Pendrak, I., Sham, K., Van Aller, G. S., Konstantinidis, A. K., Lonsdale, J. T., Janson, C. A., Qiu, X., Brandt, M., Khandekar, S. S., Silverman, C., and Head, M. S. (2003) First X-ray cocrystal structure of a bacterial FabH condensing enzyme and a small molecule inhibitor achieved using rational design and homology modeling. *J Med Chem* **46**, 5-8
31. Musayev, F., Sachdeva, S., Scarsdale, J. N., Reynolds, K. A., and Wright, H. T. (2005) Crystal structure of a substrate complex of *Mycobacterium tuberculosis* beta-ketoacyl-acyl carrier protein synthase III (FabH) with lauroyl-coenzyme A. *J Mol Biol* **346**, 1313-1321
32. Sachdeva, S., Musayev, F. N., Alhamadsheh, M. M., Scarsdale, J. N., Wright, H. T., and Reynolds, K. A. (2008) Separate entrance and exit portals for ligand traffic in *Mycobacterium tuberculosis* FabH. *Chem Biol* **15**, 402-412

33. Brown, A. K., Sridharan, S., Kremer, L., Lindenberg, S., Dover, L. G., Sacchettini, J. C., and Besra, G. S. (2005) Probing the mechanism of the Mycobacterium tuberculosis beta-ketoacyl-acyl carrier protein synthase III mtFabH: factors influencing catalysis and substrate specificity. *J Biol Chem* **280**, 32539-32547
34. Mayer, M., and Meyer, B. (2001) Group epitope mapping by saturation transfer difference NMR to identify segments of a ligand in direct contact with a protein receptor. *J Am Chem Soc* **123**, 6108-6117
35. Meyer, B., Klein, J., Mayer, M., Meinecke, R., Möller, H., Neffe, A., Schuster, O., Wülfken, J., Ding, Y., Knaie, O., Labbe, J., Palcic, M. M., Hindsgaul, O., Wagner, B., and Ernst, B. (2004) Saturation transfer difference NMR spectroscopy for identifying ligand epitopes and binding specificities. *Ernst Schering Res Found Workshop*, 149-167
36. Nikolau, B. J., Perera, M. A., Brachova, L., and Shanks, B. (2008) Platform biochemicals for a biorenewable chemical industry. *Plant J* **54**, 536-545
37. Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J., and Higgins, D. G. (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947-2948
38. Arnold, K., Bordoli, L., Kopp, J., and Schwede, T. (2006) The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics* **22**, 195-201
39. Bordoli, L., Kiefer, F., Arnold, K., Benkert, P., Battey, J., and Schwede, T. (2009) Protein structure homology modeling using SWISS-MODEL workspace. *Nat Protoc* **4**, 1-13
40. Schwede, T., Kopp, J., Guex, N., and Peitsch, M. C. (2003) SWISS-MODEL: An automated protein homology-modeling server. *Nucleic Acids Res* **31**, 3381-3385
41. Eisenberg, D., Lüthy, R., and Bowie, J. U. (1997) VERIFY3D: assessment of protein models with three-dimensional profiles. *Methods Enzymol* **277**, 396-404
42. Melo, F., and Feytmans, E. (1998) Assessing protein structures with a non-local atomic interaction energy. *J Mol Biol* **277**, 1141-1152
43. The PyMOL Molecular Graphics System, Version 1.5.0.4 Schrödinger, LLC.
44. National BioResource Project (NIG, Japan) : *E.coli*.

CHAPTER III

DELINEATING THE STRUCTURE-FUNCTION RELATIONSHIPS OF β -KETOACYL-ACP SYNTHASE III BASED ON PHYLOGENETIC AND FUNCTIONAL COMPARISONS

A paper to be submitted to the Journal of biochemistry

Shivani Garg^{1,6}, Charles Stewart^{2,7}, Marna D. Yandeau-Nelson^{3,6}, Joseph P. Noel^{4,7}, Basil J. Nikolau^{5,6}

¹Primary author and researcher, ²Collaborator, ³Co-advisor, ⁴Co-advisor, ⁵Co-advisor and author for correspondence,

⁶Center for Biorenewable Chemicals, Department of Biochemistry, Biophysics and Molecular Biology, Iowa State University, Ames, Iowa; ⁷Howard Hughes Medical Institute, The Salk Institute for Biological Studies, La Jolla, California

Abstract

β -Ketoacyl-ACP Synthase III (KASIII) initiates type II fatty acid synthesis in bacteria and plants, and exhibit diverse substrate specificity in terms of their acyl-CoA substrates. Despite the availability of several crystal structures of KASIII from various organisms, the underlying molecular basis for the differences in substrate specificity remains unclear. Recent studies have hypothesized that a specific conserved Phe residue that lines the active site pocket and the nature of two residues behind this residue (corresponding to Phe304, Val215 and Leu220 respectively in *E. coli* KASIII) are responsible for defining the substrate specificity of KASIII enzymes. In this study, we used this hypothesis to predict substrate specificities of ten putative KASIII enzymes chosen from three bacterial sources that can produce >60% branched chain fatty acids, namely *Capnocytophaga gingivalis*, *Legionella pneumophila* and *Myxococcus xanthus*, and tested these predictions, and thus the underlying hypothesis via *in vitro* activity assays and fluorescence-based thermal shift assays using different acyl-CoA substrates. Of the ten selected enzymes, we identified four that exhibited KASIII catalytic activity, and our data support the

hypothesis that the nature of the two residues behind the critical Phe governs the substrate specificity of the KASIII enzymes. Several of the selected KASIIIs were paralogs from the same bacterial strain but sorted into different sequence-based clades of the KASIII phylogenetic tree. We therefore investigated functional differences between these paralogs and developed sequence-function correlations for their divergence. We used these correlations to categorize KASIII enzymes into four structure-function groups, and identified common features in the conserved loops L3 and L9 of each group that may be responsible for the different substrate specificities exhibited by these groups. In conclusion, we have developed a structure-function framework based on functionality of previously characterized KASIII enzymes, and this framework should be useful in predicting catalytic activities based on sequence and structural models.

Introduction

β -Ketoacyl-ACP Synthase III (FabH) initiates type II fatty acid synthesis in bacteria and plants via Claisen condensation of an acyl-CoA starter molecule (e.g. acetyl-CoA) and an extender molecule (e.g. malonyl-ACP) to produce a β -ketoacyl-ACP intermediate (1). This intermediate is fully reduced to acyl-ACP and undergoes iterative chain extension to produce, in most organisms, 16- and 18-carbon fatty acids. KASIII enzymes exhibit vast diversity in substrate specificities. For example, in Gram-negative bacteria such as *Escherichia coli*, KASIII is highly selective for the straight-chain acyl-CoA (acetyl-CoA) and is therefore considered to have narrow substrate specificity (2-5). In contrast, in the Gram-positive bacteria such as *Bacillus subtilis* and *Staphylococcus aureus*, KASIII can utilize both straight-chain and branched-chain acyl-CoAs (e.g., isobutyryl-CoA and isovaleryl-CoA), and thus have broad substrate specificities (6-8). Because KASIII functionality varies so widely among organisms,

KASIII is thought to determine the fatty acid profile of an organism (9). For example, in *E. coli*, KASIII is highly selective for straight chain acyl-CoA, therefore it leads to production of straight chain fatty acids (10); whereas in *B. subtilis* and *S. aureus*, KASIII preference for branched chain substrates results in mainly branched chain fatty acid synthesis with small amounts of straight chain fatty acids (10). Therefore, depending on the nature of the starter acyl-CoA molecules (straight chain, branched chain), KASIII is capable of incorporating different functionalities into the terminal ω -end (omega-end) of the growing fatty acid chain. KASIII belongs to the thiolase family of condensing enzymes (1) and has been characterized functionally and structurally from several bacterial (2,6,7,11-20), protozoan (21,22) and plant (23-29) sources. Of all characterized KASIII, the KASIII from *E. coli* is the most extensively studied and well characterized and has been crystallized both in the presence and absence of acetyl-CoA and CoA(7,12-14,16). In addition to *E. coli*, KASIII crystal structures are available from eight Gram-negative bacteria, including *Haemophilus influenzae*, *Aquifex aeolicus* and several others (7,18). Structures are also available for KASIII from three Gram-positive bacteria (*Staphylococcus aureus* (7,8) and others (7,30)) as well as from the Gram-variable bacterium, *Mycobacterium tuberculosis*, that makes mycolic acids and its KASIII has high preference for lauroyl, palmitoyl-CoA, (31-34). However, the structure-function relationship for KASIII is not yet understood.

In various attempts to understand the underlying basis of KASIII substrate diversity (7,30), KASIII structural and sequence information have been mined and several structural motifs and residues have been proposed to govern KASIII functionality. A recent study identified that the rotamer conformation of a conserved Phe (corresponding to Phe304 of *E. coli* KASIII) is different between KASIII exhibiting narrow and broad substrate specificities. It has therefore been proposed that the rotamer conformation of the conserved Phe affects the shape of

the active site cleft and thereby affects the substrate preference for KASIII (7). Superposition of KASIII crystal structures revealed that this conserved Phe is oriented away from the active site (i.e. *anti* rotamer conformation) in KASIIIs that exhibit narrow substrate specificity (e.g. *E. coli* KASIII), whereas the Phe faces towards the active site (i.e. *gauche* rotamer conformation) in KASIIIs that exhibit broad substrate specificity (e.g. *S. aureus* KASIII). Two residues corresponding to Val215 and Leu220 in *E. coli* KASIII reside on the α -helix next to the β -sheet containing the conserved Phe and are predicted to orient the Phe in a particular rotamer conformation (*anti* vs. *gauche*). KASIIIs with broad substrate specificities typically have a Phe/Trp (corresponding to Val215 of *E. coli* KASIII) and a Met/Ile (corresponding to Leu220 of *E. coli* KASIII). These bulkier residues (Phe/Trp and Met/Ile) appear to create steric hindrance so that Phe is forced to orient toward the active site, adopting a *gauche* conformation. The nature of these two residues and the associated Phe rotamer conformations from the KASIIIs that have been characterized to date are consistent with the above stated hypothesis. However, validity of this hypothesis for previously uncharacterized KASIIIs remains to be established.

In the current study, we have applied this hypothesis to predict the functionalities of ten uncharacterized putative KASIII enzymes derived from three bacterial strains that produce predominantly branched chain fatty acids (*Capnocytophaga gingivalis*, *Legionella pneumophila*, and *Myxococcus xanthus*) (35). Using a combination of computational, biochemical and biophysical approaches, we have identified functional KASIII enzymes from each of the three bacterial strains and have experimentally validated that the conserved Phe residue (corresponding to Phe304 in *E. coli* KASIII) and the two residues behind it (corresponding to Val215 and Leu220 in *E. coli* KASIII) define KASIII substrate specificity. In addition, we have classified KASIIIs into four structure-function groups and have developed a sequence-function

correlation based on which functionality and substrate specificities of uncharacterized KASIII enzymes can be predicted. We have also demonstrated that KASIIIs from the same organism have can exhibit diverse substrate specificities and this functional divergence could explain the complex array of fatty acids produced by these organisms.

Materials and Methods

Gene cloning

The cDNAs for cgKASIIIa, cgKASIIIb, and cgKASIIIc (NCBI reference sequences - WP_002669693, WP_002667978, WP_002669050, respectively) from *Capnocytophaga gingivalis* ATCC 33624; lpKASIIIa, lpKASIIIb, lpKASIIIc, and lpKASIIId from *Legionella pneumophila str. Paris* (NCBI reference sequences – YP_123672.1, YP_124492, YP_124497, YP_123920 respectively); and mxKASIIIa, mxKASIIIb and mxKASIIIc from *Myxococcus xanthus DK 1622* (NCBI reference sequences – YP_628497.1, YP_635461.1, YP_629114.1 respectively) were codon-optimized for expression in *E. coli*, synthesized and cloned into the pUC57 vector by GenScript USA Inc. (Piscataway, NJ, USA). The synthesized genes were subsequently cloned into the pDEST-17 vector using Gateway Cloning (Invitrogen, Carlsbad, CA) according to manufacturer instructions, to generate the pDEST_cgKASIIIa, pDEST_cgKASIIIb, pDEST_cgKASIIIc, pDEST_lpKASIIIa, pDEST_lpKASIIIb, pDEST_lpKASIIIc, pDEST_lpKASIIId, pDEST_mxKASIIIa, pDEST_mxKASIIIb, and pDEST_mxKASIIIc constructs.

Four pCA24N expression vectors, each encoding an N-terminal His tag and harboring *fabD* (encodes Malonyl CoA: ACP Transacylase; MCAT) , *fabG* (encodes β -Ketoacyl ACP Reductase; KAR), *acpP* (encodes acyl carrier protein; ACP) and *aas* (encodes fused 2-

acylglycerophospho-ethanolamine acyl transferase/acyl-acyl carrier protein synthetase) respectively were obtained from the National BioResource Project (NIG, Japan) (36). The *acpP* and *aas* sequences were subcloned into the pETDUET expression vector (Novagen, Merck KGaA, Darmstadt, Germany) to generate pDUET_*acpP*_aas construct.

Expression and purification of recombinant KASIII, FabD, FabG and holo-ACP proteins.

The OverExpressTM C41 (Lucigen, Middletown, WI) *E. coli* strain was used for expression of each KASIII protein from its respective pDEST-17 construct and for the expression of MCAT (FabD), KAR (FabG) from pCA24N expression vectors and holo-ACP from pDUET expression vector. The transformants were grown at 37°C in 4L Luria-Bertani medium, which was supplemented with 100 µg/ml ampicillin (Research Products International Corps., Mount Prospect, IL). The cultures were induced by addition of IPTG (Gold Biotechnology, Olivette, MO) to a final concentration of 0.4 mM when the OD₆₀₀ was between 0.6-0.8. After incubation for 16-18 h at 25°C, cells were harvested by centrifugation (10,000 X g, 4°C, 10 min). Soluble proteins were extracted by first suspending the cell pellet in lysis buffer (0.5 M NaCl, 5 mM imidazole, 20 mM Tris-HCl pH 8.0, 0.1 mg/ml phenylmethylsulfonyl fluoride, 0.1% Triton-X 100), followed by sonication (10 s pulses separated by 3 s intervals for a total of 3 min) and centrifugation (10,000 X g, 4°C, 30 min). The soluble protein fraction was filtered through a 0.45 µ filter (Corning, the Netherlands) and applied to 8 ml Ni-NTA His-bind resin. After washing the unbound protein in buffer (0.5M NaCl, 20 mM Tris-HCl, pH 8.0) supplemented with either 20 mM (Wash Buffer I) or 40 mM imidazole (Wash Buffer II), the proteins of interest were eluted with the same buffer containing 250 mM imidazole (Elution Buffer). The purified His-tagged KASIII proteins were concentrated and exchanged into sodium phosphate buffer, pH 7.2 at 4°C with Amicon centrifugal devices (Millipore, Billerica, MA).

Protein purity was assessed by Coomassie-staining of SDS-PAGE gels, which showed presence of near-homogenous, pure proteins (greater than 95% purity). Protein concentrations were determined by Bradford assay (BioRad, Hercules, CA). The concentrated proteins were either stored at -80°C or used immediately for biochemical assays.

Spectrophotometric assay to determine KASIII activity with different substrates.

Activity of purified recombinant KASIII enzymes incubated with different acyl-CoA substrates (acetyl-CoA, propionyl-CoA, isobutyryl-CoA, isovaleryl-CoA (Sigma-Aldrich)) was ascertained using a coupled enzymatic assay. The assay was performed in 96-well plate-format with three technical replicates for each reaction condition. In a total volume of 100 μ l per reaction, the reaction mix contained 100 μ M holo-ACP, 200 μ M malonyl-CoA, 10mM DTT, 50 μ M acyl-CoA substrate and 200 μ M NADPH in 0.1 M sodium phosphate buffer (pH 7.2). This reaction mix was pre-incubated with 60 μ g of MCAT (FabD) for two minutes, to initiate synthesis of malonyl-ACP. The reaction was started by the addition of 30 μ g of KAR (FabG) and varying concentrations of KASIII enzyme (0.5-15 μ g). As KASIII condensed acyl CoA and malonyl-ACP to form 3-ketoacyl-ACP, KAR (FabG) reduced this ketoacyl-ACP intermediate to 3-hydroxy-acyl-ACP, with the concomitant oxidation of NADPH to NADP⁺. Change in absorbance of NADPH was recorded at 340 nm using a Synergy 2 Multi-Mode Microplate Reader (BioTek, Winooskit, VT). Specific activity was calculated by ascertaining the moles of NADP⁺ formed per unit time per μ g of KASIII.

Circular Dichroism (CD) spectroscopy of KASIII proteins.

All CD spectra of purified KASIII proteins (0.1–0.25 mg/ml in 10 mM sodium phosphate buffer, pH 7.2) were collected using a Jasco J-710 Spectropolarimeter, in a 0.1 cm cell at 25°C.

Far-UV spectra were recorded with a bandwidth of 1.0 nm and a time response of 8 s with total of 2 accumulations of data.

Thermal Shift assays of KASIII proteins incubated with different acyl CoA substrates.

Thermal shift assays (37) were performed with a Light Cycler 480 System (Roche Applied System) using 20 μ l reactions in a 96-well plate format. KASIII protein (2 μ M-20 μ M) was mixed with SYPRO Orange dye (5X – 10X) in 0.1 M sodium phosphate buffer, pH 7.2. Various acyl CoA ligands (coenzyme-A, acetyl-CoA, propionyl-CoA, butyryl-CoA, hexanoyl-CoA, isobutyryl-CoA, isovaleryl-CoA, 3-hydroxybutyryl-CoA, malonyl-CoA, and methylmalonyl-CoA) were added in 50-fold excess of the KASIII protein being tested. For negative controls, water was used instead of an acyl-CoA ligand. Plates were sealed with an optical sealing tape, and then were heated in the Light Cycler 480 instrument from 20 °C to 85 °C at the rate of x C/min. Melting temperatures for proteins incubated with different ligands were calculated using the LightCycler 480 Protein Melt program (Roche Applied Science, Penzberg, Germany).

Results

Sequence alignment-based predictions of substrate specificities of selected KASIII

We rationally selected ten putative KASIII enzymes (cgKASIIIa, cgKASIIIb, cgKASIIIc from *C. gingivalis*; lpKASIIIa, lpKASIIIb, lpKASIIIC, lpKASIIId from *L. pneumophila*; and mxKASIIIa, mxKASIIIb, mxKASIIIC from *M. xanthus*) from a phylogenetic analysis of KASIIIs from diverse sources (38), based on two criteria: 1) the host bacterium is Gram-negative and makes 60-85% branched chain fatty acids (BCFA) (35) because so far KASIII that have been shown to have broad substrate specificities (i.e., can use both straight and branched chain acyl

CoA primers) belong to Gram-positive bacteria that predominantly synthesize BCFA such as *B. subtilis* (6), *S. aureus* (8), *S. pneumoniae* (15), *S. glaucescens* (11) and *L. monogenes* (20), and we wanted to determine if KASIII from Gram-negative bacteria with high BCFA content display broad substrate specificities too, and 2) the putative KASIII homologs selected from a single bacterial strain must sort into different phylogenetic sub-families (as defined in (38)), which would suggest that these homologs may be structurally and functionally distinct (*figure 1*). The ten selected putative KASIIIs were analyzed for presence of conserved residues in specific domains of the protein by aligning their primary sequences with sequences of other KASIIIs that have been either functionally or structurally characterized (*figure 2*).

The catalytic triad within the active site of KASIII is comprised of the residues, Cys, His and Asn. These three residues were strictly conserved in each of the ten candidate KASIIIs with the exception of *M. xanthus* KASIIIa, in which the His and Asn are replaced by Thr and Tyr (*figure 2*) respectively. Conservation of the catalytic triad among the nine remaining enzymes suggests that each of these putative KASIIIs belong to the thiolase family of condensing enzymes (1). In addition to the catalytic triad, KASIII enzymes also have specific residues that bind to CoA and ACP (12,13,16,39,40). These binding residues were found to be conserved in only four of the putative KASIIIs, i.e., cgKASIIIa and cgKASIIIb, lpKASIIIa and mxKASIIIC. The remaining six enzymes (cgKASIIIC, lpKASIIIB, lpKASIIIC, lpKASIIID, mxKASIIIa and mxKASIIIB) show substantial divergence in the residues that are important for CoA and ACP binding (*figure 2*). Specifically, Trp32 and Arg151 of *E. coli* KASIII are required for stacking of the adenine ring of acyl-CoA (12,13,16); however only in cgKASIIIa, cgKASIIIb, lpKASIIIa and mxKASIIIC, these residues are strictly conserved. Likewise, the residues that participate in ACP binding (39,40) are conserved in cgKASIIIa, cgKASIIIb, lpKASIIIa, mxKASIIIB and

mxKASIIIc, whereas the remaining five KASIII enzymes have only 1-2 residues strictly conserved out of five residues considered crucial for ACP binding (*figure 2*).

In addition to the catalytic site residues and the CoA and ACP binding residues, we also examined the conservation of residues proposed to affect KASIII substrate specificity. In a recent study, it was proposed that the rotamer conformation of a specific Phe residue (corresponding to Phe304 in *E. coli* KASIII, and we refer to it as the ‘swinging’ Phe) determines the shape of the active site cleft, thereby defining KASIII substrate specificity (7). It was proposed that the nature of two residues behind this conserved Phe residue (corresponding to Val215 and Leu220 in *E. coli* KASIII) forces the swinging Phe to adopt either an *anti* conformation (facing away from the active site) or a *gauche* conformation (facing toward the active site) (7). In the KASIII from *E. coli*, which exhibits narrow substrate specificity (i.e. selectivity for straight chain acyl-CoAs) the residues, Val215 and Leu220, are thought to force the swinging Phe304 to adopt an *anti*-rotamer conformation. In contrast, in KASIIIs from *S. aureus* and *B. subtilis*, which exhibit broad substrate specificities (i.e. selectivity for both straight and branched chain acyl-CoAs), a bulky residue (e.g. Phe or Trp), corresponding to Val215 of *E. coli* KASIII, is thought to force the Phe to adopt a *gauche*-rotamer conformation. In most broad substrate specificity KASIIIs that have been characterized (such as KASIII from *B. subtilis* KASIIIa, KASIIIb, *S. aureus* KASIII, *S. aureus*, *S. pneumoniae*, *L. monogenes*), there is a Met or Val corresponding to Leu220 of *E. coli* KASIII (*figure 2*), which is also thought to be crucial for determining the rotamer conformation of the swinging Phe residue.

In the multiple sequence alignment, we observed that of the ten candidate KASIIIs, three enzymes, cgKASIIIa, cgKASIIIc and lpKASIIIa, have the swinging Phe and also Phe/Trp and/or Met corresponding to Phe209 and Met214 of the *S. aureus* KASIII, which exhibits broad

substrate specificity (*figure 2*). We therefore predicted that these three KASIIIb would be able to process both straight and branched chain acyl-CoA primers. cgKASIIIb has the swinging Phe, but behind this Phe it has Asn and Phe instead of Phe and Met. The effect of these two substitutions on substrate selectivity of cgKASIIIb could not be predicted (*figure 2*). The remaining six enzymes, lpKASIIIb, lpKASIIIc, lpKASIIId and mxKASIIIa, mxKASIIIb and mxKASIIId do not have the swinging Phe residue (*figure 2*), but some of them have a Phe or a Met corresponding to Phe209 and Met214 of *S. aureus* KASIII. To date, no KASIII has been characterized that does not have the conserved swinging Phe. Therefore, we could not predict the substrate specificity of these six KASIII enzymes.

We tested our hypothesis on the substrate specificities of these ten putative KASIII enzymes using two separate *in-vitro* biochemical approaches, i.e., fluorescence-based thermal shift assays and spectrophotometric enzyme assays.

Thermal shift assays distinguished between narrow and broad substrate specificity KASIII enzymes

Each of the ten candidate KASIII enzymes was purified to near homogeneity (*figure 3*) and confirmed to be folded using circular dichroism (data not shown). To assess substrate specificities for each KASIII, substrate binding was determined via fluorescence-based thermal shift assays (or differential scanning fluorimetry), described in detail elsewhere (37,41). These assays measure the change in protein melting temperature (T_m) upon addition of a substrate, buffer, salt or other additives using hydrophobic fluorescent dye that binds to the exposed hydrophobic regions of the melting protein. Positive shift in T_m indicates binding and stabilization by the substrate, whereas negative shift in T_m suggests a destabilizing effect of the substrate. The substrates tested with each KASIII included coenzyme A, straight chain (C2:0,

C3:0, C4:0, C6:0), branched chain (i-C4:0, i-C5:0), hydroxylated (3-hydroxybutyryl-CoA) and carboxylated (malonyl-CoA, methylmalonyl-CoA) acyl-CoAs. Water was used as a control for these experiments.

For each of the candidate KASIIIs, differences were seen in their inherent T_m s in water without any ligand (*table 1*). These melting temperatures were used as the baseline T_m s to measure the shift in T_m by addition of substrates.

The cgKASIIIa enzyme was most stabilized by the branched chain substrates, isobutyryl and isovaleryl-CoA, which resulted in T_m shifts significantly different from the control and all other substrates tested (p-value < 0.05; *figure 5A*). This suggests that cgKASIIIa most tightly binds the branched chain substrates. Short chain substrates (C2:0, C3:0) had a destabilizing effect on the protein, which was statistically different (p-value < 0.05) from all other substrates. In contrast, C4:0 and C6:0 straight chain substrates stabilized the enzyme to a small extent but ΔT_m was not statistically significant with respect to the control. For cgKASIIIb, each of the straight, branched, hydroxylated or carboxylated substrates were unable to increase T_m of the enzyme relative to the control. In the presence of some substrates (e.g., butyryl-CoA; *figure 5A*), the T_m for cgKASIIIb decreased, which suggests that these particular substrates had a destabilizing effect on the protein. Interestingly, cgKASIIIc was stabilized by straight chain acyl-CoAs (C2:0, C3:0, C4:0 CoAs), branched chain acyl-CoAs (i-C4:0, i-C5:0) and methylmalonyl-CoA, which showed significant change in T_m relative to the control and other substrates. The degree of stabilization increased progressively with the length of the straight chain acyl-CoA; butyryl-CoA showed 1.5- and 1.3-fold increases in ΔT_m when compared to acetyl and propionyl-CoA, respectively. The C6:0 acyl-CoA substrate exceptionally destabilized cgKASIIIc resulting in a -20 °C shift in T_m relative to water. In summary, thermal shift assay data clearly showed that

cgKASIIIb was not stabilized by any of the substrates tested, but cgKASIIIa was stabilized preferentially by branched chain substrates and cgKASIIIC was stabilized by straight, branched chain and carboxylated substrates, suggesting divergence in function among the three enzymes.

Amongst the four KASIIIs from *L. pneumophila*, lpKASIIIa was the only enzyme that could be thermally stabilized in the presence of CoA substrates, whereas the remaining three, lpKASIIIb, lpKASIIIC and lpKASIIId, could not be stabilized by any of the substrates tested (*figure 5B*). For lpKASIIIa, straight chain (C3:0, C4:0), branched chain (i-C4:0, i-C5:0) and carboxylated acyl-CoAs (malonyl, methylmalonyl-CoA) stabilized the protein and increased the T_m by 2-4°C relative to the control (p-value < 0.05). This implies that lpKASIIIa has a broad substrate preference. In contrast, the observation that the remaining three lpKASIIIs are incapable of binding acyl-CoA substrates is consistent with the observation that these three enzymes lack some of the conserved residues important for CoA binding (*figure 2*).

In case of *M. xanthus*, mxKASIIIa was immensely destabilized by most of the CoA substrates that were tested, with the exceptions of CoA and 3-hydroxybutyryl-CoA, neither of which significantly affected the T_m of the protein (*figure 5C*). Absence of the catalytic triad in this enzyme could explain the observed destabilization by acyl CoA ligands probably because of lack of binding of the acyl moieties to the protein. In contrast to mxKASIIIa, mxKASIIIb, remained largely unaffected by each substrate (p-value > 0.05 for each substrate compared to water), suggesting that it does not have affinity for any of the substrates. This data is supported by the fact that mxKASIIIb lacks some of the conserved CoA binding residues (*figure 2*). KASIIIC from *M. xanthus* was the only enzyme that was stabilized by binding to acyl-CoA substrates, specifically straight chain acyl-CoAs (C2:0, C4:0 and C6:0) and malonyl CoA, which increased the T_m of mxKASIIIC by 2-4 °C compared to the control (p-value < 0.05).

Methylmalonyl-CoA and CoA destabilized mxKASIIIc by 4°C relative to the control. The strong destabilization effect of CoA is difficult to explain without a tertiary structure of the protein.

Functional characterization of the putative enzymes

After evaluating the substrate preferences of each of the selected KASIIIs over a wide range of substrates, we quantified their specific activities using an *in vitro* spectrophotometric enzyme activity assay. CoA substrates included even and odd straight chain (C2:0, C3:0), and even and odd iso-branched chain (i-C4:0, i-C5:0) acyl-CoAs, as these are the primers used *in vivo* for *de novo* even and odd, straight and branched chain fatty acid synthesis in bacteria (10).

Of the *C. gingivalis* KASIIIs, the cgKASIIIa enzyme was most active with the branched chain acyl-CoAs, isobutyryl-CoA and isovaleryl-CoA, and had little to no activity with the straight chain substrates, propionyl-CoA and acetyl-CoA (*table 2*). In contrast, cgKASIIIb enzyme did not show any activity with straight or branched substrates, which was consistent with thermal shift assay data in which no substrates increased its melting temperature (*figure 5A*). Interestingly, cgKASIIIc exhibited at least four-fold higher specific activity with straight chain substrates as compared to branched chain substrates. In combination, the substrate binding and enzyme activity data clearly highlighted functional differences between the three KASIIIs from *C. gingivalis*.

Amongst *L. pneumophila* KASIII enzymes, lpKASIIIa was active with both straight and branched chain acyl-CoAs, with a stronger preference for branched chain substrates. For example, lpKASIIIa was eleven-fold more active on isobutyryl-CoA compared to acetyl-CoA and three-fold more active as compared to propionyl-CoA (*table 2*). In contrast, lpKASIIIb, lpKASIIIc and lpKASIIId were not active on any of the acyl-CoA substrates that were tested. This enzyme activity data is consistent with the thermal shift data where lpKASIIIa was

stabilized the most by isobutyryl CoA and lpKASIIIb, lpKASIIIC and lpKASIIId could not be thermally stabilized by any of the acyl-CoAs used in this study. For *M. xanthus*, mxKASIIIA and mxKASIIIB did not show enzymatic activity with any of the substrates tested (*table 1*). This is consistent with the results from thermal shift assays where none of the substrates could stabilize these two proteins. Moreover, mxKASIIIA lacks the conserved catalytic triad, mxKASIIIB lacks the CoA and ACP binding residues and both mxKASIIIA and KASIIIB lack the swinging Phe. In contrast, mxKASIIIC contains the conserved catalytic triad and the CoA and ACP binding sequences, but lacks the swinging Phe. Even so, mxKASIIIC displayed enzymatic activity with both straight and branched chain substrates, with a preference for straight chain substrates. This enzymatic preference for short chains is supported by the thermal shift data, in which straight chain substrates induced a higher shift in T_m as compared to branched chain substrates.

Discussion

Structural divergence among putative KASIII enzymes ensures at least one functional KASIII enzyme

This study has identified four functional KASIII enzymes, (i.e. *C. gingivalis* KASIIIA and KASIIIC, *L. pneumophila* KASIIIA and *M. xanthus* KASIIIC), from ten putative KASIII candidates chosen from three diverse bacterial sources. Although some of these putative KASIIIs enzyme were derived from the same bacterial strain, each enzyme sorted into a different sub-family of the KASIII phylogenetic tree (38). This evolutionary divergence among the sequences suggests that the proteins may have different structural features. The current study demonstrated that four of these ten proteins are functional KASIII enzymes based on their abilities to bind and enzymatically act upon at least one acyl-CoA substrate. Although KASIIIB from *C. gingivalis*

has important conserved features typical of KASIII enzymes (such as Cys-His-Asn catalytic triad, CoA and ACP binding residues), it did not show any binding and enzymatic activity with the substrates tested in this study. The remaining five KASIIIs in this study (i.e. *L. pneumophila* KASIIb, KASIIc and KASIIId, and *M. xanthus* KASIIa and KASIIb) did not show KASIII functionality and that can be correlated to absence of at least one of the following conserved domains in their primary sequences – 1) catalytic triad (for example, in mxKASIIa), ii) or the CoA binding residues (such as in lpKASIIb, lpKASIIc, lpKASIIId, mxKASIIb), or iii) the ACP binding residues (such as in lpKASIIb, lpKASIIc, lpKASIIId and mxKASIIa).

In the case of the *C. gingivalis* KASIIIs, the two functional KASIIIs clearly exhibited functional divergence. For example, cgKASIIa displayed preference for branched chain substrates, whereas cgKASIIc showed higher specificity towards straight chain substrates over branched substrates. These varied functionalities indicate that these enzymes have diverged evolutionarily and have developed specialized functions, and at the same time presence of at least one KASIII with high specificity for branched chain substrates, i.e. cgKASIIa, explains predominance of branched chain fatty acids in *C. gingivalis* (>82% BCFA (35)). Typically, bacteria have one copy of KASIII gene with so far only one known exception to the rule, i.e. *B. subtilis*, which has two KASIII homologs, bsKASIIa and bsKASIIb, that show subtle differences in their substrate preferences (6). The bsKASIIa enzyme displays higher preference for *anteiso*-branched chain substrates whereas bsKASIIb exhibits preference for *iso*-branched chain substrates (6). Presence of two KASIII homologs in *B. subtilis* is thought to be correlated to modulation of *iso*- and *anteiso*- branched chain fatty acid content in response to temperature changes (Jin and Nikolau, unpublished data). Presence of two functionally distinct KASIII

homologs in *C. gingivalis* may serve a similar function which can be explored by gene knockout studies.

In the case of the KASIII homologs from *L. pneumophila* and *M. xanthus*, only one KASIII-like sequence from each bacterial strain has been confirmed in this study as a functional KASIII (i.e. lpKASIIIa and mxKASIIIc), both of which are active on straight and branched chain substrates with higher specificity for the former substrates. However, the remaining KASIIIs in *L. pneumophila* and *M. xanthus* seem to be non-functional with the substrates tested in this study. Sorting of these KASIIIs into different sub-families of the KASIII phylogenetic tree (38) reinforces our experimental observations that each of these KASIIIs is highly diverse functionally. Also, our observation that at least one functional KASIII from each bacterial strain is capable of processing branched chain primers is consistent with the accumulation of branched-chain fatty acids in these bacterial strains, i.e. 62% and 65% BCFA in *M. xanthus* and *L. pneumophila* respectively (35).

Classification of KASIII enzymes into four distinct structure-function groups

Phylogenetic analysis of KASIII enzymes functionally characterized in this study (i.e. cgKASIIIa, cgKASIIIc, lpKASIIIa, mxKASIIIc) and KASIIIs functionally or structurally characterized in other studies (6-8,11,12,20,30) suggest that KASIII can be classified into four different structure-function groups (groups 1a, 1b, 2 and 3) based on their primary sequence similarities (*figure 5*). Correlations can be drawn between structural and functional features of KASIII from these four groups by analyzing in detail the fatty acid profiles of their host bacteria, the structural information about the rotamer conformation of the conserved Phe and the observed functionalities towards different acyl CoA primers (*table 3*).

Upon correlating the structural information with the functional data available, it is clear that KASIII_s in group 1 exhibit high specificity for straight chain acyl-CoA primers as exemplified by *E. coli* KASIII (2,12,16) (13) and belong to bacteria that make only straight chain fatty acids, such as *E. coli*, *Y. pestis* and *H. influenza* (35). This narrow substrate specificity towards straight chain acyl CoA correlates with the presence of Val and Leu (corresponding to Val215 and Leu220 of *E. coli* KASIII) in group 1 KASIII_s. Moreover, based on the crystal structures available from this group, specifically of *E. coli* (12,13,16) and *H. influenzae* KASIII_s (7), the conserved Phe in this group adopts an *anti*-rotamer conformation. This is in sharp contrast to group 3 KASIII enzymes that belong to bacteria predominantly synthesizing branched chain fatty acids, have bulkier Phe and Met residues (corresponding to Phe209 and Met214 of *S. aureus* KASIII) and exhibit high specificity for branched chain primers. For example KASIII_s from *B. subtilis* (6), *S. aureus* (8) and *Listeria monogenes* (20) have been shown to process isobutyryl and iso-valeryl-CoA substrates. cgKASIII_a, characterized in this study, also belongs to group 3 and its high specificity for branched chain primers is consistent with our prediction based on presence of the conserved Phe and the neighboring bulky Phe and Met residues. Crystal structures available from group 3 KASIII enzymes, i.e. of *T. thermophilus* (PDB code: 1UB7), *S. aureus* KASIII_s (7), show the conserved Phe (corresponding to Phe298 in *S. aureus* KASIII) in *gauche*-rotamer conformation. So, it is reasonable to predict that the remaining KASIII enzymes in this group also have the Phe in the same *gauche*-orientation, and uncharacterized enzymes from this group such as *B. licheniformis* KASIII_a and KASIII_b can be predicted to have high specificities for branched chain primers.

While KASIII enzymes in groups 1a and 3 show strong trends in their substrate preferences and presence of either Val/Leu or Phe/Met close to the conserved Phe, KASIII from

groups 1b and 2 do not exhibit a clear pattern in their functional and structural features (*table 3*). Group 1b KASIII enzymes either have the Phe (corresponding to the Phe298 of *S. aureus* KASIII) or a Met (corresponding to the Met214 of *S. aureus* KASIII) but do not have both the residues together. Functional activity data available for lpKASIIIa (characterized in this study) and *S. pneumoniae* KASIII (spKASIII) (15) from group 1b shows that these KASIIIs can utilize both straight and branched chain substrates with a higher preference for straight chain substrate. Structural information available for *E. faecalis* KASIII (7) from this group shows that its conserved Phe is oriented in *gauche*-rotamer conformation like the KASIIIs in group 3, but functional activity data is needed to further confirm if this enzyme has higher specificity for straight chain substrates like lpKASIIIa and spKASIII. Interesting feature of *A. aeolicus* KASIII (aaKASIII; PDB code: 2EBD) from this group is that it lacks the conserved Phe which is thought to be a critical residue for defining KASIII substrate specificity. Functional characterization of aaKASIII will provide insights into how absence of the conserved Phe influences substrate specificity of KASIII.

Group 2 KASIII enzymes bear resemblance to group 3 enzymes in terms of presence of Phe and Met close to the conserved Phe (*table 3*). Functional activity data for group 2 KASIIIs shows varied functionalities of enzymes of this group. For example, mxKASIIIC, cgKASIIIC, characterized in this study, can process both straight and branched chain substrates but have higher preference for straight chain substrates. Interesting feature of mxKASIIIC is the lack of the conserved Phe but it still shows KASIII functionality. *X. oryzae* KASIII (xaKASIII), that also belongs to this group, also lacks the conserved Phe and has non-conserved residues, Ile and Ala, corresponding to Phe209 and Met214 of *S. aureus* KASIII. xaKASIII has been structurally characterized (18) and shows high sequence divergence which raises questions about its

functionality as a KASIII enzyme. Functional characterization of both mxKASIIIC and xaKASIII will be insightful as both of these proteins lack the conserved Phe. This group also contains unique KASIII enzymes that can process long chain acyl-CoA substrates, i.e. *M. tuberculosis* and *M. luteus* KASIIIs (30,32,42). *S. glaucescens* KASIII, which shows high sequence identity to *M. tuberculosis* and *M. luteus* KASIII, and also has a Thr instead of Phe87 of *E. coli* KASIII, could also be capable of utilizing long chain acyl-CoA substrates. Overall, KASIII from this group show diversity in their substrate preferences. Identification and characterization of additional KASIIIs that belong to this group will help in defining the characteristic features of the group.

Other residues affecting KASIII substrate specificity

In addition to the conserved Phe (corresponding to Phe298 of *S. aureus* KASIII) which is thought to determine the shape of the active site cleft (7), there may be other residues that differ in KASIIIs from different groups and contribute to their substrate preferences. Previously, length of loop L9 and loop-loop interactions at the KASIII dimer interface (specifically between loop L9-L9' and loop L9-L3') have been suggested to be involved in determining KASIII substrate selectivity (43). Phe87' from loop L3 and Arg196' from loop L9 of *E. coli* KASIII extend from one sub-unit of the KASIII dimer into the other sub-unit, and are thought to have a role in defining KASIII substrate specificity (12,13,16). We, therefore, compared the loops L3 and L9 of the KASIIIs from the four phylogenetic groups and observed similarities and differences among groups which can be correlated to their observed functional differences (*figure 6*). For example, loop L3 of group 1a KASIII enzymes is different at three positions (corresponding to Thr81, Thr84 and Ser89 of *E. coli* KASIII) from the rest of the groups. Groups 1b, 2 and some

members of group 3 have a conserved Pro and Asp corresponding to positions 83, 84 of *E. coli* KASIII which have Ala and Thr respectively.

The Phe87 of *E. coli* KASIII is conserved in other group 1a enzymes, in some of the group 1b and in all of the group 3 members except for cgKASIIIa. Group 2 shows divergence corresponding to the Phe87 of *E. coli* KASIII, the classical examples of this group being *M. tuberculosis*, *M. luteus* KASIIIa that have a smaller Thr enabling selectivity for longer chain acyl CoAs (31,32,34). Presence of Thr in *S. glaucescens* at this position suggests that this enzyme may also be capable of processing long chain acyl-CoAs; a proposition which needs to be validated experimentally. Saturation point mutagenesis of Phe87 in *E. coli* KASIII in a previous report showed that substitution to Ile and Val resulted in activity towards C6:0 acyl-CoA substrates *in vivo* (44). mxKASIIIc, which has Ile in place of Phe87, shows binding with hexanoyl-CoA in our thermal binding assays consistent with the results of the previous report, but cgKASIIIc, which has Leu in place of Phe87, does not show binding with hexanoyl-CoA suggesting there may be other residues hindering binding of C6:0 acyl-CoA chains.

Loop L9 also shows marked differences at specific positions. For example, group 1a members have bulky Tyr/Phe corresponding to Tyr185 of *E. coli* KASIII, whereas group 3 members have a smaller residue, Gly at this position. However, group 3 members possess a bulky Tyr corresponding to Thr190 of *E. coli* KASIII. Additionally, Arg196 of *E. coli* KASIII is conserved amongst group 1a only and is not found in other groups suggesting a role of this Arg in limiting substrate specificity of group 1a members to short straight chain-acyl CoAs. In the same way, group 3 has Asn and Arg corresponding to Ala208 and Asn210 of *E. coli* KASIII respectively.

From these observed dissimilarities in the loops L3 and L9 of the four groups of KASIII enzymes, we can infer that in addition to the conserved Phe (corresponding to the Phe304 of *E. coli* KASIII that adopts different conformations in different KASIIIs (7)), subtle differences in the residues lining L3 and L9 loops, that form the KASIII dimer interface, might also be responsible for defining the KASIII primer specificity. Strategic mutations in these loop regions may enhance our understanding of the role of specific conserved residues in these loops and can help us engineer novel KASIII enzymes with unique substrate specificities. Additionally, previously uncharacterized KASIII enzymes can be sorted into one of the four phylogenetic groups based on their sequence similarities and their functionality can be predicted using the structure-function correlations drawn in this study.

Acknowledgements

This work was supported by the NSF ERC grant for Center for Biorenewable Chemicals (CBiRC), and was also supported in part by the CBiRC SLC sponsored grant. We would like to thank undergraduate students, Morgan Becker and Josh Scaralia, for their help in protein expression and purification. We would also like to acknowledge the W.M. Keck Metabolomics lab where enzyme assays were conducted.

Table 1. Melting temperatures of each of the ten selected KASIII enzymes in water and in the absence of an acyl CoA ligand. These temperatures were used as the baseline temperatures against which the effect of adding different CoA substrates was measured.

Organism	Enzyme	Melting temperature (°C)
<i>C. gingivalis</i>	KASIIIa	68.3 ± 0.17
	KASIIIb	67.8 ± 0.03
	KASIIIc	56.8 ± 0.62
<i>L. pneumophila</i>	KASIIIa	54.0 ± 0.36
	KASIIIb	60.2 ± 0.16
	KASIIIc	62.0 ± 0.03
	KASIIId	65.8 ± 0.18
<i>M. xanthus</i>	KASIIIa	64.6 ± 0.13
	KASIIIb	54.5 ± 0.15
	KASIIIc	61.4 ± 0.5

Table 2. Specific activity (nmol/min/mg) of KASIII enzymes from *C. gingivalis*, *L. pneumophila*, *M. xanthus* with straight chain (acetyl, propionyl) and branched chain (isobutyryl, isovaleryl) acyl-CoA substrates.

Organism	Enzyme	Acetyl-CoA	Propionyl-CoA	Isobutyryl-CoA	Isovaleryl-CoA
<i>C. gingivalis</i>	KASIIIa	Not detectable	25 ± 2.7	166.08 ± 82.3	53.5 ± 6.1
	KASIIIb	Not detectable	Not detectable	Not detectable	Not detectable
	KASIIIc	102.8 ± 5.7	213 ± 70	26 ± 9.1	22 ± 8.3
<i>L. pneumophila</i>	KASIIIa	27.5 ± 9.7	83.2 ± 17	304.9 ± 25.3	90.5 ± 17.7
	KASIIIb	Not detectable	Not detectable	Not detectable	Not detectable
	KASIIIc	Not detectable	Not detectable	Not detectable	Not detectable
	KASIIId	Not detectable	Not detectable	Not detectable	Not detectable
<i>M. xanthus</i>	KASIIIa	Not detectable	Not detectable	Not detectable	Not detectable
	KASIIIb	Not detectable	Not detectable	Not detectable	Not detectable
	KASIIIc	45.2 ± 10.0	37.4 ± 12.9	19.8 ± 3.0	7.6 ± 1.8

Phylogenetic group	Organismal data		Structural data		Functional data (Specific activity)		KASIII functional class
	Bacterial source of KASIII	Branched chain fatty acid content (%) ⁹	Residues affecting substrate specificity	Rotamer conformation of the swinging Phe (from crystal or predicted structures)	Straight chain acyl CoA primers utilized	Branched chain acyl CoA primers utilized	
Group 1a	<i>Escherichia coli</i>	0	F V L	Anti ¹	C2:0, C3:0 ⁴	N.D.	Narrow
	<i>Yersinia pestis</i>	0	F V L		N.D.	N.D.	
	<i>Vibrio vulnificus</i>	2	F V L		N.D.	N.D.	
	<i>Haemophilus influenzae</i>	0	F L L	Anti ¹	N.D.	N.D.	
Group 1b	<i>Legionella pneumophila</i> KASIIIa*	65	F I M		C2:0, C3:0, C4:0	i-C4:0, i-C5:0	Broad
	<i>Capnocytophaga gingivalis</i> KASIIIb*	84	F N F		No activity shown	No activity shown	
	<i>Aquifex aeolicus</i>		M V M	Not applicable	N.D.	N.D.	Broad
	<i>Vibrio cholerae</i>	5	F R M		N.D.	N.D.	
	<i>Streptococcus pneumoniae</i>	0	F F V		C2:0, C3:0, C4:0 ³	i-C4:0, i-C5:0 ³	
	<i>Enterococcus faecalis</i>	0	F F V	Gauche ¹	N.D.	N.D.	
Group 2	<i>Xanthomonas oryzae</i>	12	I E A	Not applicable	N.D.	N.D.	Broad
	<i>Myxococcus xanthus</i> KASIIIc*	62	V F A	Not applicable	C2:0, C3:0, C4:0, C6:0	i-C4:0, i-C5:0	
	<i>Capnocytophaga gingivalis</i> KASIIIc*	84	F F T		C2:0, C3:0, C4:0	i-C4:0, i-C5:0	Broad
	<i>Streptomyces glaucescens</i>	90 ⁴	F W M		C2:0, C4:0 ⁴	i-C4:0 ⁴	
	<i>Micrococcus luteus</i>	83	F W M	Gauche ²	C12:0 ²	N.D.	Broad
	<i>Mycobacterium tuberculosis</i>	Mycolic acids	Y W M	Gauche ^{1, 11}	C8:0, C10:0, C12:0, C14:0, C16:0, C18:0 ¹⁰	N.D.	Broad
Group 3	<i>Thermus thermophilus</i>	99	F F M	Gauche ¹	N.D.	N.D.	Broad
	<i>Bacillus subtilis</i> KASIIIa	95	F F M	Gauche ⁶	C2:0, C3:0, C4:0, C5:0, C6:0, C7:0, C8:0 ⁵	i-C4:0, i-C5:0, a-C5:0 ⁵	
	<i>Bacillus licheniformis</i> KASIIIa	91	F F M		N.D.	N.D.	Broad
	<i>Listeria monocogenes</i>	93	F F M		C2:0 ⁷	i-C4:0, i-C5:0, a-C5:0 ⁷	
	<i>Staphylococcus aureus</i>	69	F F M	Gauche ¹	C2:0, C3:0, C4:0, C6:0, C12:0, C16:0 ⁸	i-C4:0, i-C5:0 ⁸	Broad
	<i>Bacillus subtilis</i> KASIIIb	95	F W V	Gauche ⁶	C2:0, C3:0, C4:0, C5:0, C6:0, C7:0 ⁵	i-C4:0, i-C5:0, a-C5:0 ⁵	Broad
	<i>Bacillus licheniformis</i> KASIIIb	91	F W V		N.D.	N.D.	Broad
	<i>Capnocytophaga gingivalis</i> KASIIIa*	84	F Y M		C3:0	i-C4:0, i-C5:0	
Non-functional KASIII	<i>Myxococcus xanthus</i> KASIIIa*	62	S F W	Not applicable	No activity shown	No activity shown	
	<i>Myxococcus xanthus</i> KASIIIb*	62	V S M	Not applicable	No activity shown	No activity shown	
	<i>Legionella pneumophila</i> KASIIIb*	65	V L Y	Not applicable	No activity shown	No activity shown	
	<i>Legionella pneumophila</i> KASIIIc*	65	A K M	Not applicable	No activity shown	No activity shown	
	<i>Legionella pneumophila</i> KASIIId*	65	T L I	Not applicable	No activity shown	No activity shown	

1- Gajiwala et al, 2009, 2 - Pereira et al, 2012, 3 - Khandekar et al, 2001, 4 - Han et al, 1998, 5 - Choi et al, 2000, 6 - Garg et al, 2013, 7 - Singh et al, 2009, 8 - Qiu et al, 2005, 9 - Ratledge and Wilkinson, 10 - Choi et al, 2000, 11 - Scarsdale et al, 2001

Table 3. Comparison of structural and functional features of KASIII identified in this study (marked with *) and other studies. Also included are some KASIIIs that are previously uncharacterized. KASIII enzymes are grouped into four groups based on the phylogenetic analysis presented in this study (figure 5). Non-functional KASIII proteins identified in this study are highlighted in gray. (N.D. – Not determined).

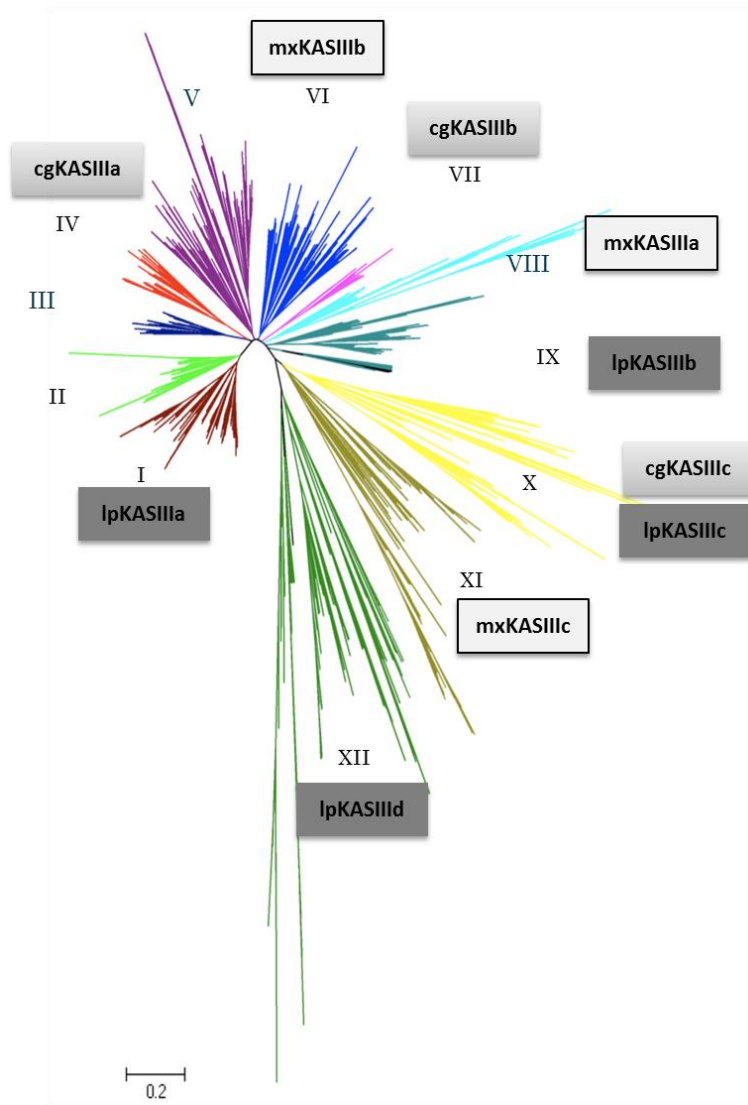


Figure 1. Putative KASIII genes chosen for functional characterization. Ten putative KASIII genes were chosen from previously computationally identified 2,308 KASIII sequences (38). We selected three KASIII homologs from *Capnocytophaga gingivalis* (cgKASIIIa, cgKASIIIb, cgKASIIIc), four KASIII homologs from *Legionella pneumophila* (lpKASIIIa, lpKASIIIb, lpKASIIIc, lpKASIIId) and three KASIII homologs from *Myxococcus xanthus* (mxKASIIIa, mxKASIIIb, mxKASIIIc) which all sort into different sub-families of the KASIII phylogenetic tree.

Bacterial source of KASIII	Branched chain fatty acid content (%)	(A)			(B)			(C)					(D)			(E)		
		Active site residues			CoA binding residues			ACP binding residues					Residues affecting active site cleft shape			Residues at dimer interface		
		1	2	2			1		2	2	2	2	3	2	2		1	1
		1	4	7	2	3	5	3	1	1	4	5	0	1	2	8	9	5
		2	4	4	8	2	1	6	3	4	9	3	4	5	0	7	6	7
<i>Escherichia coli</i>	0	C	H	N	T	W	R	R	F	K	R	A	F	V	L	F	R	F
<i>Haemophilus influenzae</i>	0	C	H	N	T	W	R	R	F	K	R	A	F	L	L	Y	R	F
<i>Yersinia pestis</i>	0	C	H	N	T	W	R	R	F	K	R	A	F	V	L	F	R	F
<i>Vibrio vulnificus</i>	2	C	H	N	T	W	R	R	F	K	R	A	F	V	L	F	R	F
<i>Vibrio cholerae</i>	5	C	H	N	T	W	R	R	F	K	R	T	F	R	M	I	R	F
<i>Capnocytophaga gingivalis</i> KASIIIa	84	C	H	N	T	W	R	R	F	K	R	A	F	Y	M	V	I	F
<i>Capnocytophaga gingivalis</i> KASIIIb	84	C	H	N	T	W	R	R	F	K	R	F	F	N	F	F	E	F
<i>Capnocytophaga gingivalis</i> KASIIIc	84	C	H	N	W	K	K	K	F	N	H	K	F	F	T	L	-	F
<i>Legionella pneumophila</i> KASIIIa	65	C	H	N	T	W	R	R	F	K	R	A	F	I	M	F	-	F
<i>Legionella pneumophila</i> KASIIIb	65	C	H	N	K	K	E	E	K	E	K	Q	V	L	Y	T	T	L
<i>Legionella pneumophila</i> KASIIIc	65	C	H	N	S	V	F	S	F	V	N	S	A	K	M	S	A	F
<i>Legionella pneumophila</i> KASIIId	65	C	H	N	L	S	L	K	F	K	L	H	T	L	I	I	-	F
<i>Myxococcus xanthus</i> KASIIIa	62	C	T	Y	K	W	K	N	V	Q	R	A	S	F	W	S	-	F
<i>Myxococcus xanthus</i> KASIIIb	62	C	H	N	W	R	Y	K	A	R	R	A	V	S	M	F	-	F
<i>Myxococcus xanthus</i> KASIIIc	62	C	H	N	T	W	R	R	Q	K	R	A	V	F	A	I	-	W
<i>Streptococcus pneumoniae</i>	0	C	H	N	T	W	R	R	F	K	R	K	F	F	V	M	F	F
<i>Enterococcus faecalis</i>	0	C	H	N	T	W	R	R	F	D	R	K	F	F	V	M	F	F
<i>Bacteroides vulgatus</i>	55	C	H	N	T	W	R	R	F	D	R	A	F	Y	M	F	Y	F
<i>Bacillus subtilis</i> KASIIIa	95	C	H	N	T	W	R	R	F	K	R	A	F	F	M	F	-	F
<i>Bacillus subtilis</i> KASIIIb	95	C	H	N	T	W	R	R	Y	K	R	S	F	W	V	F	G	F
<i>Bacillus licheniformis</i> KASIIIa	91	C	H	N	T	W	R	R	F	K	R	A	F	F	M	F	-	F
<i>Bacillus licheniformis</i> KASIIIb	91	C	H	N	T	W	R	R	Y	K	R	S	F	W	V	F	G	F
<i>Listeria monogenes</i>	93	C	H	N	T	W	R	R	F	R	R	A	F	F	M	F	-	F
<i>Micrococcus luteus</i>	83	C	H	N	S	W	R	R	F	R	R	E	F	W	M	T	D	L
<i>Staphylococcus aureus</i>	69	C	H	N	T	W	R	M	F	K	R	S	F	F	M	F	-	F

Figure 2. Alignment of conserved residues of the ten putative KASIII enzymes with KASIII enzymes from both Gram-negative and Gram-positive bacteria. Seventeen conserved residues, grouped into five categories, are aligned for experimentally confirmed and putative KASIII enzymes from Gram-negative bacteria (highlighted in blue) and Gram-positive bacteria (highlighted in green). Branched chain fatty acid content of each bacterium is listed and color coded so that darker green color shade indicates higher branched chain fatty acid content, and blue indicates very low or no branched chain fatty acids.

Figure 2 (contd.). Coloring scheme for residues: Black – residues conserved in KASIII_s from both Gram-negative and Gram-positive bacteria, Gray – homologous substitutions for the conserved residues, Blue – residues similar to those in *E. coli* KASIII but not in *S. aureus* KASIII, Green – residues similar to those in *S. aureus* KASIII but not in *E. coli* KASIII, White – non-conserved residues. Numbering scheme corresponds to the *E. coli* KASIII residues.

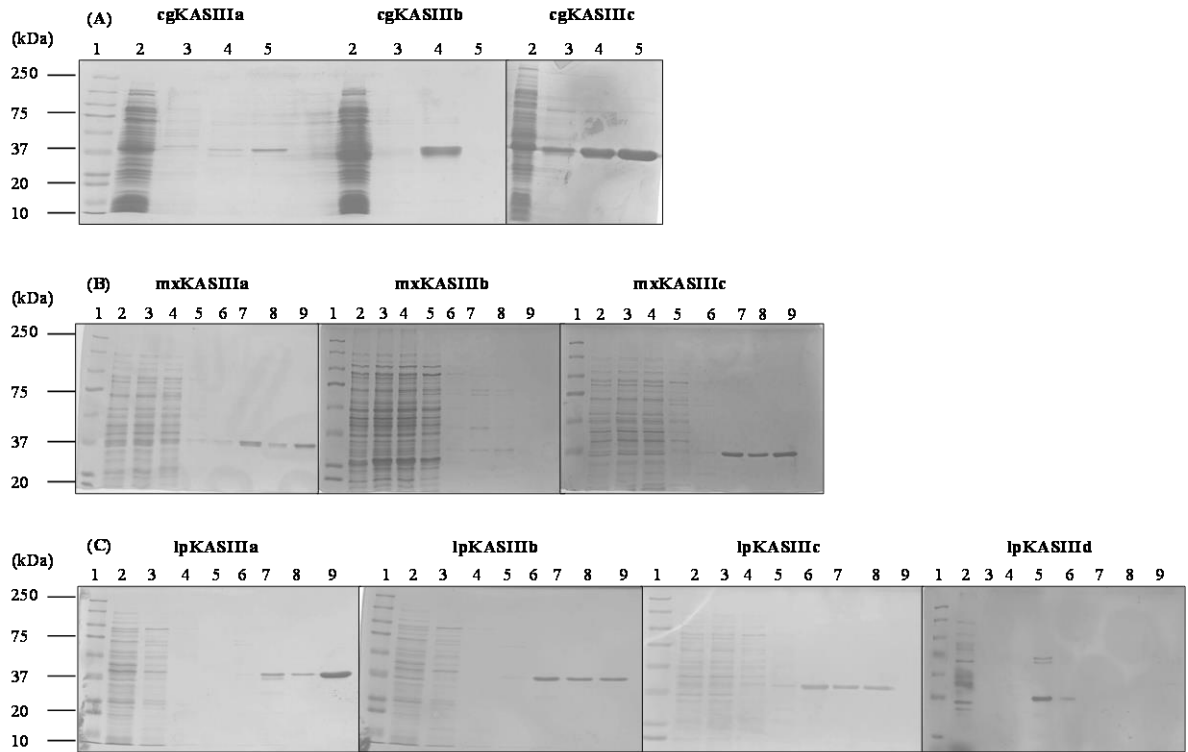


Figure 3. Expression and purification of the candidate KASIII proteins. Each of the ten putative His-tagged KASIII proteins were purified by Ni-NTA affinity chromatography and assessed for homogeneity and purity using Coomassie-stained gels. Panel (A) shows purified *C. gingivalis* KASIIIa, KASIIIb and KASIIIc (1: Protein ladder, 2: Crude extract, 3: First wash, 4: Elution with 80 mM imidazole, 5: Elution with 250 mM imidazole), Panel (B) shows purified *M. xanthus* KASIIIa, KASIIIb and KASIIIc (1: Protein ladder, 2: Crude extract, 3: Flow through, 4: First wash with 20 mM imidazole, 5: Second wash with 20 mM imidazole, 6: Wash with 40 mM imidazole buffer, 7: Elution with 80 mM imidazole, 8: Second elution with 80 mM imidazole, 9: Elution with 250 mM imidazole). Panel (C) shows purified *L. pneumophila* KASIIIa, KASIIIb, KASIIIc and KASIIId (1: Protein ladder, 2: Crude extract, 3: Flow through, 4: First wash with 20 mM imidazole, 5: Second wash with 20 mM imidazole, 6: Wash with 40 mM imidazole buffer, 7: Elution with 80 mM imidazole, 8: Second elution with 80 mM imidazole, 9: Elution with 250 mM imidazole).

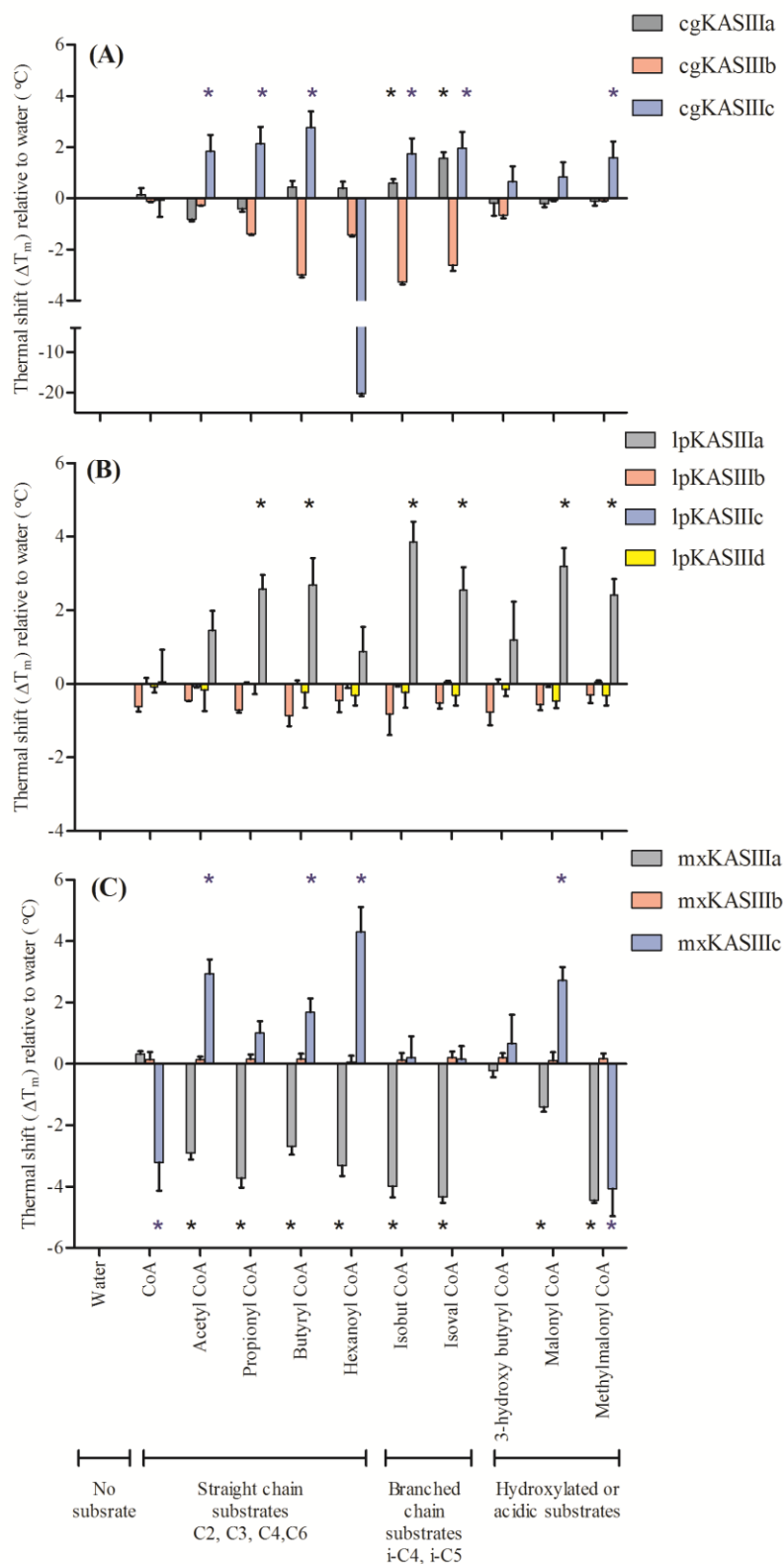


Figure 4. Comparison of substrate preferences of the ten candidate KASIII proteins via fluorescence-based thermal shift assay. Change in the melting temperature (ΔT_m) with respect to the baseline T_m in water upon adding different acyl CoA substrates was studied for selected KASIII proteins, i.e., *C. gingivalis* KASIIIa, KASIIIb and KASIIIc (A), *L. pneumophila* KASIIIa, KASIIIb, KASIIIc and KASIIId (B), and *M. xanthus* KASIIIa, KASIIIb and KASIIIc (C). Each data point represents average of four technical replicates.

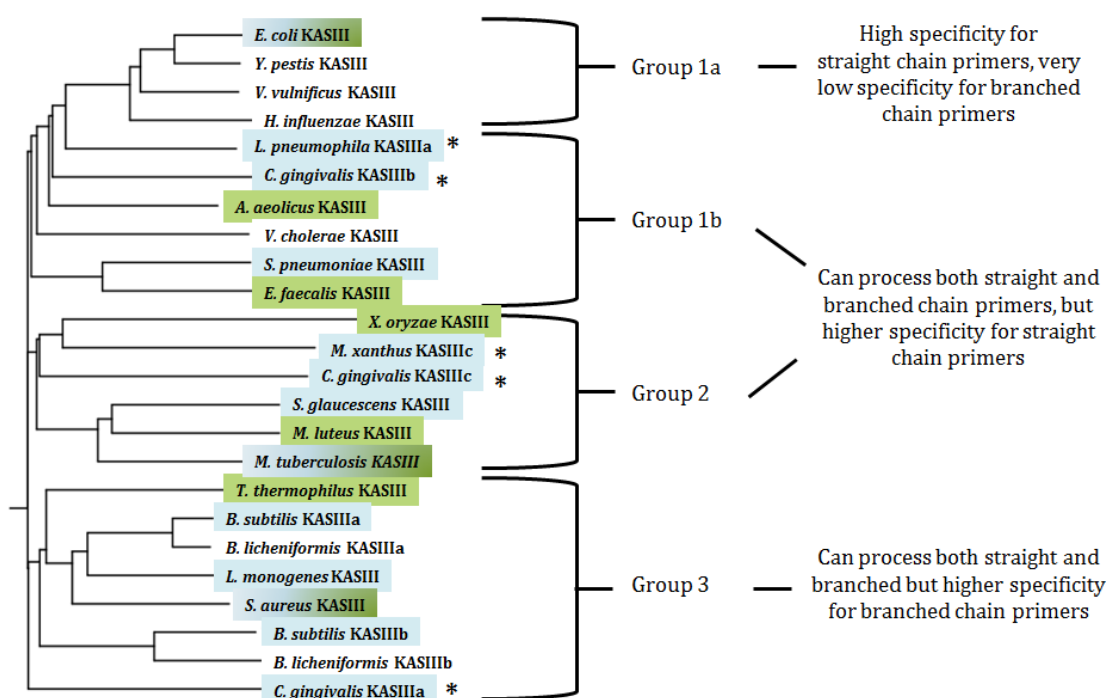


Figure 5. Classification of KASIII enzymes into four distinct structure-function groups. Phylogenetic tree was created for the functional KASIII enzymes identified in this study (marked with *) and other KASIII enzymes for which structural data is available (highlighted in green) or functional data is available (highlighted in blue) or are still uncharacterized. ClustalW2 Phylogeny program was used to create the tree.

		Loop L3				Loop L9									
		80	89	182									210		
Group 1a	<i>Escherichia coli</i>	TTSATHA	FPS	DGSYSEL	LTPNA								DRVNPEN-SIH	TMA-GN	
	<i>Yersinia pestis</i>	TSSSHA	FPS	DGRFYSEL	LAPYP								DR-QQDQ-PAY	VTMA-GN	
	<i>Vibrio vulnificus</i>	TSSGHT	FPS	DGEFSLLES	LEVPEV								ER-GKDA-DKW	THMA-GN	
	<i>Haemophilus influenzae</i>	TSSHSA	FPS	SADKNNA	VIQAQP								ER-GIEK-SGY	IEMQ-GN	
Group 1b	<i>Legionella pneumophila KASIIIA*</i>	TCTPLMF	FPS	AYDSDKL	LVIRNS								TF--EQD-RAT	TGMR-GN	
	<i>Capnocytophaga gingivalis KASIIIB*</i>	TCSPLYF	FPG	EGAHABEL	IAVTPGVG								KKWVDLIKENDPEDT-SYY	PYMN-GO	
	<i>Aquifex aeolicus</i>	TLTPQKR	FPS	EGSLEEL	LHADN								CGY--TRMK-GR		
	<i>Vibrio cholerae</i>	TCSPTSL	LEN	DAQGRDI	IAVPKEFG								TAM--DRFDADNGYWA-FDFV-GK		
	<i>Streptococcus pneumoniae</i>	TITPLSM	MPS	DGSRSECT	TYGHSGHLH								SPF--SDQESA-DSF	TKMD-GR	
	<i>Enterococcus faecalis</i>	TVTPLFN	MPS	DGQRWTA	ITSGYTINE								SPF--YQGHKQA-SKT	IQME-GR	
Group 2	<i>Xanthomonas oryzae</i>	SVSRMYL	EFS	TRSATEWNKI	CRG								N-LDR-	MVTD-TR	
	<i>Myxococcus xanthus KASIIIC*</i>	GCSFTFC	ICA	TPSGADKVR	VP								R-MGH-FTQN-CG		
	<i>Capnocytophaga gingivalis KASIIIC*</i>	TCTPHIL	LEP	DGTGAENT	LIRNGGTRS							AR--	NENPDDW-DNF	IDMK-GAI	
	<i>Streptomyces glaucescens</i>	TVTHFKQ	TSA	EGDKAET	KQTVPWTDYRE								G-GVER-FPA	TTQE-GQ	
	<i>Micrococcus luteus</i>	TVTTFPHA	TES	DGERWST	SMTHSQLELRDAVEHARTTGDA	SAIT							GA-EGML-WPT	TRQD-GPI	
	<i>Mycobacterium tuberculosis</i>	TNVHFLQ	TBP	DGEQADA	LRQDIDWITFAQ								N-PSGP-RPF	VRIE-GPI	
Group 3	<i>Thermus thermophilus</i>	TNTTHAL	FED	DGTGAKEL	YHACVA							PRL	PD-GTSM-KNR	LYMN-GR	
	<i>Bacillus subtilis KASIIIA</i>	TVTTHQS	FPP	DGTGSGHLY	YN								E-KRH-TIMN-GR		
	<i>Bacillus licheniformis KASIIIA</i>	TVTTHQA	FEP	DGRGKHLY	LD								E-KDH-TIMN-GR		
	<i>Listeria monocogenes</i>	TVTTCAT	FEP	DGSGSKYL	INID								E-NKK-LYMN-GR		
	<i>Staphylococcus aureus</i>	TATCDMP	FEP	DGTGSKHY	LD								KDTGK-LKMN-GR		
	<i>Bacillus subtilis KASIIIB</i>	TVTSTYA	FEP	SGNGSLI	YRAGLR							NEI	NGVQLVG-SGK	MVCN-GR	
	<i>Bacillus licheniformis KASIIIB</i>	TITPLIA	FPS	KGDGRHI	YRTGLR								SDL	KGEPLSG-EGK	MVCN-GR
	<i>Capnocytophaga gingivalis KASIIIA*</i>	SVTHMP	VAI	DGTGREY	IKHEAGSSI								LPTT	IETLQEG-KNN	IYQD-GK
	Non-functional KASIIIs	<i>Myxococcus xanthus KASIIIA*</i>	TCTPLYL	SEA	NGEYHDAL	CVYTGGTNR							PA	TA-E-S-	LETGGK
		<i>Myxococcus xanthus KASIIIB*</i>	TCTPLAPHNH		SGNAYGD	IVRKGGGLLK							LPYQ	PGRMRPA-DMA	FVVQ-GFI
<i>Legionella pneumophila KASIIIB*</i>		STTNKR	FTE	DGQVNAFT	TQGVPF							PT	HE-QIDS-VGY	HTV-GD	
<i>Legionella pneumophila KASIIIC*</i>		TISGYFP	SCP	DGEYDAVS	TAGGSRIPAA								VC-NDNE-QFYITIKE-ST		
<i>Legionella pneumophila KASIIID*</i>		CGAGQP	IFC	YSLGAFCEV	QAGTTLM							PPA	SEYDKK-YGL	FKMD-GK	

Figure 6. Comparison of conserved L3 and L9 loop regions of KASIII from four phylogenetic groups. Residues from loops L3 and L9 are aligned for KASIII characterized in this study and other studies. Conserved residues in each group are highlighted by a colored box around them (red – residues conserved in group 1a, blue – residues conserved in groups 1b and 2, orange – residues conserved in group 3). Strictly conserved residues among all KASIII sequences are highlighted in black; homologous substitutions are highlighted in gray.

References

1. Heath, R. J., and Rock, C. O. (2002) The Claisen condensation in biology. *Nat Prod Rep* 19, 581-596
2. Tsay, J. T., Oh, W., Larson, T. J., Jackowski, S., and Rock, C. O. (1992) Isolation and characterization of the beta-ketoacyl-acyl carrier protein synthase III gene (fabH) from *Escherichia coli* K-12. *J Biol Chem* 267, 6807-6814
3. Jackowski, S., Murphy, C. M., Cronan, J. E., and Rock, C. O. (1989) Acetoacetyl-acyl carrier protein synthase. A target for the antibiotic thiolactomycin. *J Biol Chem* 264, 7624-7629
4. Jackowski, S., and Rock, C. O. (1987) Acetoacetyl-acyl carrier protein synthase, a potential regulator of fatty acid biosynthesis in bacteria. *J Biol Chem* 262, 7927-7931
5. Heath, R. J., and Rock, C. O. (1996) Inhibition of beta-ketoacyl-acyl carrier protein synthase III (FabH) by acyl-acyl carrier protein in *Escherichia coli*. *J Biol Chem* 271, 10996-11000
6. Choi, K. H., Heath, R. J., and Rock, C. O. (2000) beta-ketoacyl-acyl carrier protein synthase III (FabH) is a determining factor in branched-chain fatty acid biosynthesis. *J Bacteriol* 182, 365-370
7. Gajiwala, K. S., Margosiak, S., Lu, J., Cortez, J., Su, Y., Nie, Z., and Appelt, K. (2009) Crystal structures of bacterial FabH suggest a molecular basis for the substrate specificity of the enzyme. *FEBS Lett* 583, 2939-2946
8. Qiu, X., Choudhry, A. E., Janson, C. A., Grooms, M., Daines, R. A., Lonsdale, J. T., and Khandekar, S. S. (2005) Crystal structure and substrate specificity of the beta-ketoacyl-acyl carrier protein synthase III (FabH) from *Staphylococcus aureus*. *Protein Sci* 14, 2087-2094
9. Lu, Y. J., Zhang, Y. M., and Rock, C. O. (2004) Product diversity and regulation of type II fatty acid synthases. *Biochem Cell Biol* 82, 145-155
10. Kaneda, T. (1991) Iso- and anteiso-fatty acids in bacteria: biosynthesis, function, and taxonomic significance. *Microbiol Rev* 55, 288-302
11. Han, L., Lobo, S., and Reynolds, K. A. (1998) Characterization of beta-ketoacyl-acyl carrier protein synthase III from *Streptomyces glaucescens* and its role in initiation of fatty acid biosynthesis. *J Bacteriol* 180, 4481-4486
12. Qiu, X., Janson, C. A., Konstantinidis, A. K., Nwagwu, S., Silverman, C., Smith, W. W., Khandekar, S., Lonsdale, J., and Abdel-Meguid, S. S. (1999) Crystal structure of beta-ketoacyl-acyl carrier protein synthase III. A key condensing enzyme in bacterial fatty acid biosynthesis. *J Biol Chem* 274, 36465-36471

13. Davies, C., Heath, R. J., White, S. W., and Rock, C. O. (2000) The 1.8 Å crystal structure and active-site architecture of beta-ketoacyl-acyl carrier protein synthase III (FabH) from *Escherichia coli*. *Structure* 8, 185-195
14. Khandekar, S. S., Konstantinidis, A. K., Silverman, C., Janson, C. A., McNulty, D. E., Nwagwu, S., Van Aller, G. S., Doyle, M. L., Kane, J. F., Qiu, X., and Lonsdale, J. (2000) Expression, purification, and crystallization of the *Escherichia coli* selenomethionyl beta-ketoacyl-acyl carrier protein synthase III. *Biochem Biophys Res Commun* 270, 100-107
15. Khandekar, S. S., Gentry, D. R., Van Aller, G. S., Warren, P., Xiang, H., Silverman, C., Doyle, M. L., Chambers, P. A., Konstantinidis, A. K., Brandt, M., Daines, R. A., and Lonsdale, J. T. (2001) Identification, substrate specificity, and inhibition of the *Streptococcus pneumoniae* beta-ketoacyl-acyl carrier protein synthase III (FabH). *J Biol Chem* 276, 30024-30030
16. Qiu, X., Janson, C. A., Smith, W. W., Head, M., Lonsdale, J., and Konstantinidis, A. K. (2001) Refined structures of beta-ketoacyl-acyl carrier protein synthase III. *J Mol Biol* 307, 341-356
17. Revill, W. P., Bibb, M. J., Scheu, A. K., Kieser, H. J., and Hopwood, D. A. (2001) Beta-ketoacyl acyl carrier protein synthase III (FabH) is essential for fatty acid biosynthesis in *Streptomyces coelicolor* A3(2). *J Bacteriol* 183, 3526-3530
18. Huynh, K. H., Natarajan, S., Song, N. H., Ngo, P. T., Ahn, Y. J., Kim, J. G., Lee, B. M., Eo, Y. D., and Kang, L. W. (2009) Cloning, expression, crystallization and preliminary X-ray crystallographic analysis of beta-ketoacyl-ACP synthase III (FabH) from *Xanthomonas oryzae* pv. *oryzae*. *Acta Crystallogr Sect F Struct Biol Cryst Commun* 65, 460-462
19. Wen, L., Chmielowski, J. N., Bohn, K. C., Huang, J. K., Timsina, Y. N., Kodali, P., and Pathak, A. K. (2009) Functional expression of *Francisella tularensis* FabH and FabI, potential antibacterial targets. *Protein Expr Purif* 65, 83-91
20. Singh, A. K., Zhang, Y. M., Zhu, K., Subramanian, C., Li, Z., Jayaswal, R. K., Gatto, C., Rock, C. O., and Wilkinson, B. J. (2009) FabH selectivity for anteiso branched-chain fatty acid precursors in low-temperature adaptation in *Listeria monocytogenes*. *FEMS Microbiol Lett* 301, 188-192
21. Waters, N. C., Kopydlowski, K. M., Guszczynski, T., Wei, L., Sellers, P., Ferlan, J. T., Lee, P. J., Li, Z., Woodard, C. L., Shallom, S., Gardner, M. J., and Prigge, S. T. (2002) Functional characterization of the acyl carrier protein (PfACP) and beta-ketoacyl ACP synthase III (PfKASIII) from *Plasmodium falciparum*. *Mol Biochem Parasitol* 123, 85-94
22. Prigge, S. T., He, X., Gerena, L., Waters, N. C., and Reynolds, K. A. (2003) The initiating steps of a type II fatty acid synthase in *Plasmodium falciparum* are catalyzed by pfACP, pfMCAT, and pfKASIII. *Biochemistry* 42, 1160-1169

23. Clough, R. C., Matthis, A. L., Barnum, S. R., and Jaworski, J. G. (1992) Purification and characterization of 3-ketoacyl-acyl carrier protein synthase III from spinach. A condensing enzyme utilizing acetyl-coenzyme A to initiate fatty acid synthesis. *J Biol Chem* 267, 20992-20998
24. Jaworski, J. G., Clough, R. C., and Barnum, S. R. (1989) A Cerulenin Insensitive Short Chain 3-Ketoacyl-Acyl Carrier Protein Synthase in *Spinacia oleracea* Leaves. *Plant Physiol* 90, 41-44
25. Tai, H., Post-Beittenmiller, D., and Jaworski, J. G. (1994) Cloning of a cDNA encoding 3-ketoacyl-acyl carrier protein synthase III from *Arabidopsis*. *Plant Physiol* 106, 801-802
26. Abbadi, A., Brummel, M., Schütt, B. S., Slabaugh, M. B., Schuch, R., and Spener, F. (2000) Reaction mechanism of recombinant 3-oxoacyl-(acyl-carrier-protein) synthase III from *Cuphea wrightii* embryo, a fatty acid synthase type II condensing enzyme. *Biochem J* 345 Pt 1, 153-160
27. Dehesh, K., Tai, H., Edwards, P., Byrne, J., and Jaworski, J. G. (2001) Overexpression of 3-ketoacyl-acyl-carrier protein synthase IIIs in plants reduces the rate of lipid synthesis. *Plant Physiol* 125, 1103-1114
28. Li, J., Li, M. R., Wu, P. Z., Tian, C. E., Jiang, H. W., and Wu, G. J. (2008) Molecular cloning and expression analysis of a gene encoding a putative beta-ketoacyl-acyl carrier protein (ACP) synthase III (KAS III) from *Jatropha curcas*. *Tree Physiol* 28, 921-927
29. González-Mellado, D., von Wettstein-Knowles, P., Garcés, R., and Martínez-Force, E. (2010) The role of beta-ketoacyl-acyl carrier protein synthase III in the condensation steps of fatty acid biosynthesis in sunflower. *Planta* 231, 1277-1289
30. Pereira, J. H., Goh, E. B., Keasling, J. D., Beller, H. R., and Adams, P. D. (2012) Structure of FabH and factors affecting the distribution of branched fatty acids in *Micrococcus luteus*. *Acta Crystallogr D Biol Crystallogr* 68, 1320-1328
31. Scarsdale, J. N., Kazanina, G., He, X., Reynolds, K. A., and Wright, H. T. (2001) Crystal structure of the *Mycobacterium tuberculosis* beta-ketoacyl-acyl carrier protein synthase III. *J Biol Chem* 276, 20516-20522
32. Musayev, F., Sachdeva, S., Scarsdale, J. N., Reynolds, K. A., and Wright, H. T. (2005) Crystal structure of a substrate complex of *Mycobacterium tuberculosis* beta-ketoacyl-acyl carrier protein synthase III (FabH) with lauroyl-coenzyme A. *J Mol Biol* 346, 1313-1321
33. Sachdeva, S., Musayev, F. N., Alhamadsheh, M. M., Scarsdale, J. N., Wright, H. T., and Reynolds, K. A. (2008) Separate entrance and exit portals for ligand traffic in *Mycobacterium tuberculosis* FabH. *Chem Biol* 15, 402-412

34. Brown, A. K., Sridharan, S., Kremer, L., Lindenberg, S., Dover, L. G., Sacchettini, J. C., and Besra, G. S. (2005) Probing the mechanism of the Mycobacterium tuberculosis beta-ketoacyl-acyl carrier protein synthase III mtFabH: factors influencing catalysis and substrate specificity. *J Biol Chem* 280, 32539-32547
35. Ratledge, C., and Wilkinson, S. G. (1988) *Microbial Lipids*, Academic Press, UK
36. National BioResource Project (NIG, Japan) : E.coli.
37. Niesen, F. H., Berglund, H., and Vedadi, M. (2007) The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nat Protoc* 2, 2212-2221
38. Chen, Y., Kelly, E. E., Masluk, R. P., Nelson, C. L., Cantu, D. C., and Reilly, P. J. (2011) Structural classification and properties of ketoacyl synthases. *Protein Sci* 20, 1659-1667
39. Misra, N., Patra, M. C., Panda, P. K., Sukla, L. B., and Mishra, B. K. (2013) Homology modeling and docking studies of FabH (β -ketoacyl-ACP synthase III) enzyme involved in type II fatty acid biosynthesis of *Chlorella variabilis*: a potential algal feedstock for biofuel production. *J Biomol Struct Dyn* 31, 241-257
40. Zhang, Y. M., Rao, M. S., Heath, R. J., Price, A. C., Olson, A. J., Rock, C. O., and White, S. W. (2001) Identification and analysis of the acyl carrier protein (ACP) docking site on beta-ketoacyl-ACP synthase III. *J Biol Chem* 276, 8231-8238
41. Lo, M. C., Aulabaugh, A., Jin, G., Cowling, R., Bard, J., Malamas, M., and Ellestad, G. (2004) Evaluation of fluorescence-based thermal shift assays for hit identification in drug discovery. *Anal Biochem* 332, 153-159
42. Choi, K. H., Kremer, L., Besra, G. S., and Rock, C. O. (2000) Identification and substrate specificity of beta -ketoacyl (acyl carrier protein) synthase III (mtFabH) from *Mycobacterium tuberculosis*. *J Biol Chem* 275, 28201-28207
43. Pan, H., Tsai, S., Meadows, E. S., Miercke, L. J., Keatinge-Clay, A. T., O'Connell, J., Khosla, C., and Stroud, R. M. (2002) Crystal structure of the priming beta-ketosynthase from the R1128 polyketide biosynthetic pathway. *Structure* 10, 1559-1568
44. Nomura, C. T., Taguchi, K., Taguchi, S., and Doi, Y. (2004) Coexpression of genetically engineered 3-ketoacyl-ACP synthase III (fabH) and polyhydroxyalkanoate synthase (phaC) genes leads to short-chain-length-medium-chain-length polyhydroxyalkanoate copolymer production from glucose in *Escherichia coli* JM109. *Appl Environ Microbiol* 70, 999-1007

CHAPTER IV

IDENTIFICATION OF KASIII ENZYMES WITH NOVEL SUBSTRATE SPECIFICITIES: DEMONSTRATION OF *IN VIVO* PRODUCTION OF NOVEL ω -1 HYDROXYLATED FATTY ACIDS USING A NOVEL KASIII

A paper to be submitted to the Journal of Biological Chemistry

Shivani Garg^{1,7}, Huanan Jin^{2,7}, Charles Stewart^{3,8}, Marna D. Yandean-Nelson^{4,7}, Joseph P. Noel^{5,8}, Basil J. Nikolau^{6,7}

¹Primary author and researcher, ²Co-author and researcher, ³Collaborator, ⁴Co-advisor, ⁵Co-advisor, ⁶Co-advisor and author for correspondence

⁷Center for Biorenewable Chemicals, Department of Biochemistry, Biophysics and Molecular Biology, Iowa State University, Ames, Iowa; ⁸Howard Hughes Medical Institute, The Salk Institute for Biological Studies, La Jolla, California

Abstract

Most organisms typically synthesize straight chain fatty acids. A few organisms however produce unusual fatty acids that carry chemical functionalities near the ω -end of the molecule, for example ω -cyclic, ω -phenylic, ω -branched, or ω -halogenated fatty acids. There are two enzymatic determinants that control this ω -group functionality, the availability of appropriate acyl-CoA primer for fatty acid synthase, and the substrate specificity of the enzyme that initiates fatty acid synthesis, 3-ketoacyl-ACP synthase III (KASIII). In this study, we have biochemically and biophysically characterized unique KASIII enzymes from four diverse bacterial sources, *C. gingivalis*, *B. subtilis*, *T. aquaticus* and *A. acidocaldarius*. From this set, we have identified several KASIIIs that possess novel substrate specificities towards atypical fatty acid synthesis primers, including hydroxylated (3-hydroxybutyryl-CoA), unsaturated (crotonyl-CoA) and aromatic (benzoyl-CoA) primers. Additionally, we have demonstrated *in-vivo* production of ω -1 hydroxy-fatty acids in a modified strain of *R. rubrum*, by recombinantly expressing the *A. acidocaldarius* KASIII, which utilizes 3-

hydroxybutyryl-CoA. Identification of such unique KASIII enzymes introduces the opportunity to synthesize diverse ω -functionalized fatty acids, which can serve as chemical intermediates for production of bio-based chemicals.

Introduction

Fatty acids form an integral component of bacterial and plant cell membranes, and their synthesis and nature (straight chain or branched chain) is regulated by a highly functionally diverse enzyme, β -ketoacyl acyl carrier protein synthase III (KASIII) (1). KASIII initiates type II fatty acid synthesis in bacteria and plants, which involves dissociable enzyme units that carry out various steps of fatty acid synthesis. Specifically, KASIII catalyzes Claisen condensation by condensing a starter molecule (an acyl-CoA), typically acetyl-CoA in case of *Escherichia coli*, with an extender molecule, malonyl-ACP, to give a β -ketoacyl-ACP intermediate that undergoes a series of iterative reactions leading to saturated fatty acids that are usually 14-18 carbon long (2).

KASIII has been characterized from many bacterial (3-17), plant (18-25) and protozoan (26,27) sources and several of these characterized KASIII enzymes have been crystallized (3-5,8,10,11,13,17,28-32). *E. coli* KASII is the most extensively characterized KASIII, which has been crystallized both in the presence (3,4) and absence of ligands (8,11,28), and also with an indole analogue inhibitor (33). Functionally, KASIII exhibits vast diversity because it can utilize a variety of substrates. For example, KASIII has been shown to utilize straight chain primers (e.g. acetyl-CoA, propionyl-CoA) in Gram-negative bacteria like *E. coli*, which leads to the production of straight chain fatty acids (6,9,34-36). Several Gram-positive bacteria, such as *Bacillus subtilis* and *Staphylococcus aureus*, have been

shown to utilize medium chain and branched chain primers (e.g. butyryl-CoA, isobutyryl-CoA, isovaleryl-CoA), resulting in both straight and branched chain fatty acids (5,9,28). The KASIII from *Mycobacterium tuberculosis* is unique in that it can utilize long chain acyl-CoA primers (e.g. lauroyl-CoA), and therefore produce very long chain fatty acids, such as mycolic acids (10,29,30,32). These collective works demonstrate that KASIII enzymes have the ability to incorporate different functionalities into the omega end of the precursor fatty acid and thereby can diversify the products of fatty acid pathway.

Although KASIII enzymes from different organisms have been shown to be active with straight, branched and long chain acyl-CoA primers (5,7,9,12,15,35-38), specificity toward other acyl-CoAs (e.g. hydroxyl, aromatic, cyclic) have yet to be explored. The presence of unusual fatty acids within some plants and bacteria suggest that KASIIIs within these organisms are capable of utilizing unusual CoAs. For example, ω -fluorinated fatty acids are found in the toxic plant, *Dichapetalum toxicarium* (39) and seed oils from Araceae *sp.* (40) are rich in phenyl-terminated fatty acids. The thermophilic bacterium, *Alicyclobacillus acidocaldarius* naturally produces ω -alicyclic fatty acids (41-43) and other bacterial strains, including *Geobacillus stearothermophilus*, *Meiothermus ruber* (44) and *Bacillus subtilis* (45) have been shown to produce ω -alicyclic fatty acids upon feeding with appropriate starter ω -cyclic acids (cyclohexane, cycloheptane). Taken together, these observations suggest that the KASIIIs in these organisms are capable of utilizing unusual CoA thioesters that have cyclic, fluorinated or phenylic terminal groups. It is tenable to hypothesize that such KASIIIs may even be able to accept substrates that do not occur naturally in bacteria, such as hydroxylated, aromatic or unsaturated acyl-CoAs.

Using both *in vivo* and *in vitro* characterization methods, the current study, identified four KASIII enzymes that utilize unusual substrates, specifically hydroxylated, unsaturated and aromatic acyl-CoAs. A KASIII with specificity toward 3-hydroxy butyryl-CoA was selected to demonstrate *in vivo* production of novel ω -functionalized fatty acids by co-expressing one of the KASIIIs with enzymes that synthesize its specific substrate. The KASIII was co-expressed with the *phaA* and *phaB* genes from *Rhodospirillum rubrum* (46), which make hydroxylated acyl-CoA precursors, and the resulting strain produced omega-branched-hydroxy fatty acids. This study, therefore, unveils the new biocatalytic potential of KASIII enzymes and provides proof of concept for the synthesis of omega-functionalized fatty acids. Such novel functionalized fatty acids may have myriad applications in a bio-based chemical industry (47), and could help in replacement of petroleum based feedstock for producing commercial chemicals.

Materials and Methods

Construction of a *B. subtilis* *fabH* deletion mutant

The genes *yjaX*, which encodes bsKASIIIa, and *yhfB*, which encodes bsKASIIIb in *B. subtilis*, were deleted using derivatives of the pMUTIN4 vector, pMU4A and pMU4B, respectively (48). The pMU4A vector contains two *yjaX*-derived fragments, one upstream and one downstream of the *fabH* open reading frame (ORF). The upstream portion (897 bp in total) spanned from 860-bp upstream of the *bfabHA* ORF to 37-bp within this ORF. The downstream fragment spanned from 799-bp downstream to 98-bp within the 3'-end of the *bfabHA* ORF. These two fragments were initially PCR amplified from *B. subtilis* str. 168 ((*Bacillus* Genetic Stock Center, Columbus, Ohio)) with the primer pairs (Table 1), AUf-PacI and AUr-SalI, and ADf-SalI and ADr-PstI, and both fragments were cloned into pMUTIN4,

at the PacI and PstI sites, respectively. The resulting pMU4A vector contains an in-frame 135 bp-*bfabHA* fragment that was missing 804-bp from the middle of the ORF; the fact that this deletion allele carries an in-frame ORF avoids any polar effect on the downstream genes of the *bfabHA*-containing transcription unit. The vectors pMU4B and pUCB-erm, which were used to generate deletion allele for *bfabHB*, were constructed by an analogous procedure, except that the vector pUCB-erm was constructed from plasmid pUC19 and the erythromycin-resistant gene *erm* was inserted between the downstream and upstream DNA fragments of the *bfabHB* ORF.

Two single deletion strains of *B. subtilis*, each lacking either *bfabHA* or *bfabHB*, were generated by homologous recombination via a two-step procedure using the vectors, pMU4A and pMU4B, respectively. Briefly, plasmid pMU4A or pMU4B was transformed into the wild-type strain *B. subtilis str.* 168, followed by selection for erythromycin-resistance. The recovered integrant colonies were grown in LB liquid medium without erythromycin, and the overnight cultures were diluted 1:10⁷, and 100 µl of the diluted culture was plated on LB medium with IPTG and X-gal. Resultant white colonies indicated the loss of the *lacZ*-containing pMUTIN4 sequence, which would result in the deletion of the appropriate KASIII-coding gene. Deletion mutants were confirmed via PCR amplification across the deleted portion of *either bfabHA or bfabHB*.

The double deletion mutant strain, *ΔbfabHA ΔbfabHB::erm* was generated by homologous recombination via a one-step procedure. The *bfabHB*-deletion plasmid, pUCB-erm was linearized via digestion with EcoRI and subsequently transformed into the mutant *ΔbfabHA* strain. Resultant *ΔbfabHA ΔbfabHB* double mutant colonies were selected on media containing erythromycin and anteiso-C16:0 fatty acid. PCR confirmation of the

ΔbfabHB and *ΔbfabHB::erm* alleles were performed using the primer pairs (Table 1), lofAf/lofAr and lofBf /lofBr, respectively.

Complementation of the *B. subtilis* fabH deletion mutant with recombinant KASIII

The pUCB-erm-fabH vector was constructed from plasmid pUC19. The erythromycin-resistant gene, *erm* and the exogenous fabH gene (a KASIII-encoding gene from *C. gingivalis*, *A. acidocaldarius* or *T. aquaticus*) under the control of the promoter Pspac, were inserted between the downstream and upstream DNA fragments of the 135-bp ORF in the *ΔbfabHB* allele of the *ΔbfabHA ΔbfabHB* double mutant. Complementation of the *B. subtilis* fabH deletion mutant with the recombinant KASIII was checked by growing the resultant *B. subtilis* construct in absence of branched chain fatty acids.

Fatty acid analysis bacterial strains containing recombinant KASIII genes

A 0.5 ml aliquot of an overnight *B. subtilis* culture normalized for equal cell density was collected by centrifugation at 13,200×g for 30 seconds. The cell pellet was suspended in 50 ml minimal medium or LB rich medium. Cell cultures were shaken at 250 rpm. Cells were collected at late log phase by centrifugation at 5,000×g for 10 minutes. The collected cell pellets were lyophilized and stored at -20 °C until analysis. Lipids were extracted from lyophilized bacterial cell pellets using chloroform/methanol (49) and fatty acids were then converted to picolinyl esters (lipidlibrary.aocs.org/ms/ms02/index.htm) (50) or methyl esters using methanolic-HCl at 80 °C for 60 minutes (51,52). The recovered picolinyl esters or fatty acid methyl esters were concentrated under a stream of nitrogen gas and analyzed with GC-MS interfaced with a Mass Detector 5973 (Agilent Technologies, Santa Clara, CA). The double bond positions in unsaturated fatty acids were determined by GC-MS analysis of dimethyl disulfide adducts (53).

Gene cloning

The cDNAs for cgKASIIIa, cgKASIIIb, cgKASIIIc (NCBI reference sequences - WP_002669693, WP_002667978, WP_002669050, respectively) from *Capnocytophaga gingivalis* ATCC 33624; aaKASIII from *Alicyclobacillus acidocaldarius* subsp. *acidocaldarius* DSM 446 (NCBI reference sequence – YP_003183476.1); and taKASIII from *Thermus aquaticus* (NCBI reference sequence – WP_003048115.1) were codon-optimized for expression in *E. coli*, synthesized and cloned into the pUC57 vector by GenScript USA Inc. (Piscataway, NJ, USA). The synthesized genes were subsequently cloned into the pDEST-17 vector via Gateway Cloning (Invitrogen, Carlsbad, CA) according to manufacturer instructions to generate pDEST_cgKASIIIa, pDEST_cgKASIIIb, pDEST_cgKASIIIc, pDEST_aaKASIII and pDEST_taKASIII constructs. Each pDEST17 construct encoded an N-terminal His-tag. The *E. coli fabH* gene was PCR-amplified from *E. coli* strain MG1655 (*E. coli* Genetic Stock Center, New Haven, CT), and cloned into pDEST17 vector using Gateway cloning (Invitrogen, Carlsbad, CA), resulting in the pDEST_ecKASIII construct. The *yhfb* gene encoding *B. subtilis* KASIIIb was PCR-amplified from *B. subtilis* strain 168 and cloned into the pDEST17 expression vector via Gateway cloning methods to generate pDEST_bsKASIIIb. The resulting plasmids were confirmed by sequencing.

pCA24N expression vectors were obtained from the National BioResource Project (NIG, Japan) (54) that harbor the *fabD*, *fabG*, *acpP* and *aas* genes, which encode Malonyl-CoA: ACP Transacylase (MCAT or FabD), β -Ketoacyl ACP Reductase (KAR or FabG), holo-ACP and fused 2-acylglycerophospho-ethanolamine acyl transferase/acyl-acyl carrier protein synthetase, respectively. The *acpP* and *aas* genes were subsequently cloned together

into the pETDUET vector (Novagen, Merck KGaA, Darmstadt, Germany). Each of these constructs encodes an N-terminal His-tag.

Expression and purification of recombinant KASIIIs, FabD, FabG and holo-ACP proteins

The OverExpressTM C41 (Lucigen, Middletown, WI) strain was used for expression of all KASIII proteins from their respective pDEST17 constructs. The transformants were grown at 37°C in 4L Luria-Bertani medium, which was supplemented with 100 µg/ml ampicillin (Research Products International Corps., Mount Prospect, IL). The cultures were induced by addition of IPTG (Gold Biotechnology, Olivette, MO) to a final concentration of 0.4 mM when the OD₆₀₀ was between 0.6-0.8. After incubation for 16-18 h at 25°C, cells were harvested by centrifugation (10,000 X g, 4°C, 10 min). Soluble proteins were extracted by first suspending the cell pellet in lysis buffer (0.5 M NaCl, 5 mM imidazole, 20 mM Tris-HCl, pH 8.0, 0.1 mg/ml phenylmethylsulfonyl fluoride, 0.1% Triton-X 100), followed by sonication (10 s pulses separated by 3 s intervals for a total of 3 min) and centrifugation (10,000 X g, 4°C, 30 min). The soluble protein fraction was filtered through a 0.45 µ filter (Corning, the Netherlands) and applied to 8 ml Ni-NTA His-bind resin. After washing the unbound protein in buffer (0.5M NaCl, 20 mM Tris-HCl, pH 8.0) supplemented with either 20 mM (Wash Buffer I) or 40 mM imidazole (Wash Buffer II), the proteins of interest were eluted with the same buffer containing 250 mM imidazole (Elution Buffer). The purified His-tagged KASIII proteins were dialyzed against 0.1 M sodium phosphate buffer, pH 7.2 and concentrated using 10,000 molecular weight cut-off ultrafiltration centrifugal filters (Millipore, Billerica, MA) at 4°C. The concentrated proteins were stored at -80°C or immediately used for spectrophotometric activity assay and thermal shift assays. Protein

purity was assessed by Coomassie-staining SDS-PAGE gels, which showed presence of near-homogenous, pure proteins (greater than 95% purity). Protein concentrations were determined by Bradford's assay (BioRad, Hercules, CA). The concentrated proteins were either stored at -80°C or immediately used for spectrophotometric activity assays and thermal shift assays.

Also, the three recombinant proteins (FabD, FabG and holo-ACP encoding for Malonyl-CoA: ACP Transacylase, β -Ketoacyl ACP reductase and holo-ACP) with N-terminal His-tag were purified to near-homogeneity using the same procedure as described for the purification for KASIII proteins. Purity of these proteins was assessed by running an SDS-PAGE gel.

Spectrophotometric assay to determine KASIII activity with different substrates

The enzymatic activities of the purified recombinant KASIII enzymes were ascertained via a coupled enzymatic assay. Six acyl-CoA substrates were used in this assay, acetyl-CoA, hexanoyl-CoA, isobutyryl-CoA, 3-hydroxybutyryl-CoA, crotonyl-CoA and benzoyl-CoA (Sigma-Aldrich). The assay was performed in 96-well plate format with three technical replicates for each reaction condition. In a total volume of 100 μ l for each reaction, the reaction mix containing 100 μ M holo-ACP, 200 μ M malonyl-CoA, 10 mM DTT, 50 μ M acyl-CoA substrate and 200 μ M NADPH in 0.1 M sodium phosphate buffer (pH 7.2) was pre-incubated with 60 μ g of MCAT (FabD) for two minutes to initiate synthesis of malonyl-ACP. The reaction was started by the addition of 30 μ g of KAR (FabG) and varying concentrations of KASIII enzyme (0.5-15 μ g). As KASIII condensed acyl-CoA and malonyl-ACP to form 3-ketoacyl-ACP, KAR (FabG) reduced this ketoacyl-ACP intermediate to 3-hydroxy-acyl ACP in presence of NADPH. Change in absorbance of

NADPH, as it was being consumed by KAR and was being converted to NADP⁺, was recorded at 340 nm using Synergy 2 Multi-Mode Microplate Reader (BioTek, Winooskit, VT). Specific activity was calculated by ascertaining the moles of product (NADP⁺) formed per unit time per µg of KASIII.

Circular Dichroism (CD) spectroscopy of KASIII proteins

All CD spectra of purified KASIII proteins (0.1–0.25 mg/ml in 10 mM sodium phosphate buffer, pH 7.2) were collected using a Jasco J-710 Spectropolarimeter, in a 0.1cm cell at 25°C. Far-UV spectra were recorded with a bandwidth of 1.0 nm and a time response of 8 s with a total of 2 accumulations of data.

Thermal Shift assays of KASIII proteins incubated with different acyl-CoA substrates

Thermal shift assays (55,56) were performed with a Light Cycler 480 System (Roche Applied System) using 20 µl reactions in a 96-well plate format (56). KASIII protein (2 µM–20 µM) was mixed with SYPRO Orange dye (5X – 10X) in 0.1 M sodium phosphate buffer, pH 7.2. Various acyl-CoA ligands (coenzyme-A, acetyl-CoA, propionyl-CoA, butyryl-CoA, hexanoyl-CoA, isobutyryl-CoA, isovaleryl-CoA, 3-hydroxybutyryl-CoA, malonyl-CoA, methylmalonyl-CoA, crotonyl-CoA, benzoyl-CoA and phenylacetyl-CoA) were added in 50-fold excess of the KASIII protein being tested. For negative controls, water was used instead of an acyl-CoA ligand. Plates were sealed with an optical sealing tape, and then were heated in the Light Cycler 480 instrument from 20 °C to 95 °C at the rate of 1 °C/min. Melting temperatures for proteins incubated with different ligands were calculated using the LightCycler 480 Protein Melt program (Roche Applied Science, Penzberg, Germany).

Cloning of *A. acidocaldarius* KASIII into *Rhodospirillum rubrum*, and co-expression with phaA, phaB genes

The *A. acidocaldarius* KASIII gene was cloned into the *phaC2* locus (Aru_2413) in the *R. rubrum* genome via a double crossover recombination event. The *R. rubrum* recipient strain for this experiment was the *phaC* triple mutant ($\Delta phaC1\Delta phaC2\Delta phaC3$) that lacked any PhaC activity (46). First, the upstream flanking sequence (922bp) of the *R. rubrum* *phaC2* gene was cloned upstream of the *A. acidocaldarius* KASIII sequence, and this chimeric construct was introduced into the *E. coli* strain, S17-1. Transconjugation was induced by overnight co-incubation of the *R. rubrum* *phaC* triple mutant ($\Delta phaC1\Delta phaC2\Delta phaC3$) with the S17-1 strain harboring the chimeric construct on a 0.22 μ m filter. The bacterial mixture was subsequently cultured on minimal medium containing 25 μ g/ml gentamicin for one to two weeks. The resulting colonies carried the product of a single recombination crossover event, which integrates the *A. acidocaldarius* KASIII gene at the *phaC2* gene (Aru_2413) locus in the *R. rubrum* $\Delta phaC1\Delta phaC2\Delta phaC3$ strain. These colonies were streaked out on another minimal medium containing 25 μ g/ml gentamicin for colony purification. The resulting colonies were cultured in SMN rich medium for two to three days in the light without gentamicin selection. Finally, double-crossover events were identified via selection on SMN rich medium containing 5% sucrose. Resultant $\Delta phaC1\Delta phaC2\Delta phaC3$ *aaKASIII* strains were confirmed via PCR amplification.

Results

Strategic selection of KASIII genes from diverse bacterial sources

Six KASIII genes were selected from bacteria that produce high percentages of either branched chain fatty acids (either *iso*- or *anteiso*-), or ω -cyclic fatty acids, based on the hypothesis that the KASIII enzymes from such bacteria will have large enough substrate

binding pockets that can accommodate other bulky substrates such as aromatic or hydroxylated or unsaturated acyl-CoAs. Specifically, we selected a KASIII gene from the acidothermophile, *A. acidocaldarius*, which synthesizes 59% ω -alicyclic fatty acids and 36% branched chain fatty acids (57), and a KASIII gene from the thermophile, *T. aquaticus*, which produces 95% branched chain fatty acids (57). Three KASIII genes (cgKASIIIa, cgKASIIIb, cgKASIIIc) were selected from the bacterium *C. gingivalis*, which make 84% branched chain fatty acids (57). Finally, the KASIIIb gene from *B. subtilis* was selected based on previous reports that ω -alicyclic fatty acids could be produced by a *B. subtilis* strain that was fed precursor alicyclic carboxylic acids (e.g. cyclobutanecarboxylic acid, cyclohexanecarboxylic acid) (45). The KASIII from *E. coli* (ecKASIII) was selected as a control enzyme based on its highly selectivity for acetyl-CoA and its inability to process branched-chain substrates (9,34-36,58).

Whereas the three KASIII from *C. gingivalis* (cgKASIIIa, cgKASIIIb, cgKASIIIc) and the *B. subtilis* KASIIIb (bsKASIIIb) have been functionally characterized ((9); Chapters 2 and 3 of this dissertation) with branched and straight-chain substrates, the KASIII proteins from *A. acidocaldarius* (aaKASIII) and *T. aquaticus* (taKASIII) have not yet been characterized. Bioinformatics analyses of the primary structures of aaKASIII and taKASIII revealed that each possesses the catalytic triad composed of Cys, His and Asn (*figure 1*), suggesting that both belong to the family of decarboxylating thiolase enzymes (2).

A comparative analysis of aaKASIII and taKASIII sequences with other functionally characterized KASIII enzymes (*figure 1*) showed that each has the conserved CoA and ACP binding residues. Both aaKASIII and taKASIII also have the conserved Phe (*figure 1*), which is thought to adopt different rotamer conformations in different KASIII enzymes and

thereby determine substrate specificities (17,28). The two neighboring residues behind this conserved Phe are Leu222 and Met227 in aaKASIII, and Phe306 and Met222 in taKASIII, which correspond to the bulky residues that are typically found in KASIIIs that have broad substrate specificities (e.g. Phe298 and Met303 in *S. aureus* KASIII) (*figure 1*). Based on the conservation of the conserved Phe (*anti*-rotamer or *gauche*-rotamer (nomenclature defined in Chapter 2 of this dissertation), and presence of either Phe or Met behind it, aaKASIII and taKASIII both belong to can be predicted to have broad substrate specificities (Chapter 3 of this dissertation). As described in Chapter 3, bsKASIIIb, cgKASIIIa and cgKASIIIc also have broad substrate specificities whereas cgKASIIIb does not exhibit any functional activity.

***In vivo* analysis of KASIII function in the $\Delta fabH$ *B. subtilis* strain**

The bacterium *B. subtilis* primarily synthesizes branched chain fatty acids, and the two KASIII homologs, *yjaX* (bsKASIIIa) and *yhfB* (bsKASIIIb), have been shown to possess high specificity for branched chain primers (9). Deletion of these two endogenous KASIII genes results in a lethal phenotype (10), which can be rescued by growing the $\Delta yjaX \Delta yhfB$ double mutant in the presence of branched-chain fatty acids. To assess whether the five selected KASIII proteins can catalyze the production of branched chain fatty acids, each KASIII was integrated into the genome of the $\Delta \Delta yjaX \Delta \Delta yhfB$ double mutant, as described in the Methods. KASIII expression in these $\Delta yjaX \Delta yhfB$ KASIII strains was induced by IPTG and grown in the absence of exogenously supplied branched-chain fatty acids. Thus, we used a *B. subtilis fabH* deletion mutant system to screen for KASIII enzymes that have propensity for branched chain acyl-CoA primers and can synthesize branched chain fatty acids, and therefore, are capable of rescuing the lethal *fabH* deletion phenotype. Of the five KASIII

genes tested (i.e. cgKASIIIa, cgKASIIIb, cgKASIIIc, taKASIII and aaKASIII) in the $\Delta yjaX \Delta yhfB$ double mutant, only aaKASIII, taKASIII and cgKASIIIa could rescue the lethal phenotype, suggesting that these KASIII enzymes can initiate synthesis of branched chain fatty acids.

Fatty acid analysis of the $\Delta yjaX \Delta yhfB$ strains harboring aaKASIII, taKASIII or cgKASIIIa revealed that the KASIII from *A. acidocaldarius* could produce 48% *anteiso*-fatty acids and 30% odd number *iso*-fatty acids when expressed in the *B. subtilis* $\Delta yjaX \Delta yhfB$ strain (table 2). *T. aquaticus* KASIII could make the highest percentage of *anteiso*-fatty acids, i.e. 52%, and synthesized 27% odd number *iso*-fatty acids (table 2). *C. gingivalis* KASIIIa synthesized the least amount of *anteiso*-fatty acids (27%), but maximum amount of odd number *iso*-fatty acids (59%) (table 2). These data collectively suggest that aaKASIII and taKASIII have preferences for *anteiso* branched acyl-CoA primers (which could be anteisovaleryl-CoA), whereas cgKASIIIa is more selective for *iso*-acyl-CoA primers (which could be isobutyryl-CoA or isovaleryl-CoA).

The inability of cgKASIIIb and cgKASIIIc to rescue the lethal *B. subtilis fabH* deletion mutant implies that either these KASIII enzymes cannot utilize branched chain acyl-CoA primers *in vivo* or they do not utilize the branched chain primers at a level sufficient to rescue the lethal phenotype observed for the *B. subtilis* $\Delta yjaX \Delta yhfB$ strain.

Identification of KASIII enzymes that can process atypical substrates

The three enzymes (aaKASIII, taKASIII and cgKASIIIa) that rescued the lethal phenotype of the *B. subtilis fabH* deletion mutant were purified to homogeneity (figure 2) and then screened for binding with both typical and atypical KASIII substrates. Typical substrates included straight and branched chain acyl-CoA primers, whereas atypical substrates included diacid (malonyl-CoA, methylmalonyl-CoA), hydroxylated (3-

hydroxybutyryl-CoA), unsaturated (crotonyl-CoA), and aromatic (such as benzoyl- and phenylacetyl-CoA) acyl-CoAs (*figure 3*). As standards, we included *E. coli* KASIII, which is a well-characterized KASIII enzyme that is highly selective for straight chain acyl-CoAs, and *B. subtilis* KASIIIb, which is highly selective for branched chain acyl-CoAs. We predicted that *B. subtilis* KASIIIb would be more receptive to atypical substrates than *E. coli* KASIII because of the former enzyme's propensity to process branched chain substrates, it should have a bigger substrate binding pocket as compared to *E. coli* KASIII allowing binding by bulkier substrates such as hydroxylated, unsaturated or aromatic acyl-CoAs.

The ability of KASIIIs to bind to these substrates was measured via a fluorescence-based thermal shift assay (55,56,59), which measures the thermal stability of a protein in the presence and absence of a specific ligand or substrate. A positive shift in melting temperature of the protein (T_m) in the presence of a substrate is correlated with substrate binding and concomitant stabilization of the enzyme, whereas a negative shift in T_m suggests destabilization of the protein by the substrate. Thermal shift analysis showed that the KASIIIs from *B. subtilis*, *A. acidocaldarius* and *T. aquaticus* bound a broad range of substrates, whereas *C. gingivalis* KASIIIa and *E. coli* KASIII bound to a relatively narrow range of substrates.

Inherent T_{ms} in water without any ligand were measured for each of the KASIIIs (*table 3*). These melting temperatures were used as the baseline T_{ms} to measure the shift in T_m by addition of substrate. An interesting feature that we noted in the thermal shift assays is the unusual thermal stability of *A. acidocaldarius* and *T. aquaticus* KASIII enzymes that had baseline T_{ms} of $\sim 73^\circ\text{C}$ and $\sim 84^\circ\text{C}$ respectively (*table 3*). These melting temperatures are

considerably higher than those for other KASIII_s studied in this report and can be correlated to the ability of these two bacteria to survive at extremely high temperatures.

Consistent with previous reports that *E. coli* KASIII has narrow substrate specificity, ecKASIII could bind only short, straight chain acyl-CoAs, (acetyl- and propionyl-CoA) (*figure 3*). In contrast, *C. gingivalis* KASIII_a mainly bound branched chain (i.e. isovaleryl- and isobutyryl-CoAs) and unsaturated (crotonyl-CoA; T_m shift of 2°C) substrates that resulted in statistically significant thermal shifts (*figure 3*).

B. subtilis KASIII_b and *A. acidocaldarius* KASIII exhibited the broadest ranges of substrate specificities. In particular, *B. subtilis* KASIII_b could bind the straight chain substrates (propionyl-CoA and butyryl-CoA), the branched isobutyryl- and isovaleryl-CoAs, a diacidic substrate (methylmalonyl-CoA), an unsaturated substrate (crotonyl-CoA) and aromatic substrate (phenylacetyl-CoA). Each of these substrates induced at least a 6°C increase in the T_m of bsKASIII_b (*figure 3*). In contrast some substrates (e.g. hexanoyl-CoA and benzoyl-CoA) apparently destabilized the protein, as evidenced by a decrease in the T_m of bsKASIII_b. The binding capacity of *A. acidocaldarius* KASIII was equally broad. aaKASIII bound to the straight chain acetyl-CoA and hexanoyl-CoA substrates, the branched isobutyryl- and isovaleryl-CoAs, a hydroxylated substrate (3-hydroxybutyryl-CoA), an unsaturated substrate (crotonyl-CoA) and the aromatic substrate (benzoyl-CoA), each of which resulted in at least a 4°C increase in the T_m of the enzyme. Malonyl-CoA and phenylacetyl-CoA also bound to aaKASIII, inducing a 2°C increase in T_m .

T. aquaticus KASIII bound with rather a small range of substrates, specifically straight butyryl- and hexanoyl-CoA, and branched isobutyryl- and isovaleryl-CoA substrates that resulted in 2-6°C increase in T_m as compared to the baseline T_m .

Demonstration of *in vivo* production of omega-branched hydroxy fatty acids using a novel KASIII

The KASIII from *A. acidocaldarius* binds to 3-hydroxybutyryl-CoA (*figure 3*), suggesting that aaKASIII could produce 3-hydroxy fatty acids within a bacterial strain that generated the 3-hydroxybutyryl-CoA. Recently, a set of three PHA polymerase genes (*phaC1*, *phaC2*, and *phaC3*) was shown to be necessary for the production of polyhydroxyalkanoates (PHAs) in the bacterium, *Rhodospirillum rubrum* (46). The triple *phaC* mutant (Δ *phaC1* Δ *phaC2* Δ *phaC3*) is incapable of accumulating PHA polymer, and shows only a slight impact on growth characteristics. This strain, therefore, has the capacity to generate 3-hydroxybutyryl-CoA via *phaA* mediated condensation of two acetyl-CoA molecules to give acetoacetyl-CoA which is then reduced by *phaB* to give 3-hydroxybutyryl-CoA (*figure 5*). This substrate could potentially be utilized by a KASIII to produce ω -1-hydroxy-fatty acids. We tested this hypothesis by recombinantly expressing the *A. acidocaldarius* KASIII in an *R. rubrum* strain that lacks functional *phaC* enzymes (Δ *phaC1* Δ *phaC2* Δ *phaC3* triple mutant) and then analyzing the fatty acids produced.

Our data showed that the substrate 3-hydroxybutyryl-CoA in *R. rubrum* was being utilized by *A. acidocaldarius* KASIII and was metabolized via fatty acid synthesis pathway resulting in C15:0 ω -1 hydroxy fatty acids (*figure 4*). This result demonstrated *in-vivo* production of bi-functional fatty acids in a recombinant bacterial host.

Discussion

Typical substrates for KASIII enzymes are either linear chain acyl-CoAs (for example, acetyl-CoA and propionyl-CoA in case of *E. coli* KASIII (9,60)) or branched chain acyl-CoAs (for example, isobutyryl-CoA and isovaleryl-CoA in case of *B. subtilis* and *S.*

aureus KASIIIs (5,9)). Substrate specificity for *M. tuberculosis* KASIII is unique as it can utilize long chain acyl-CoA, i.e. lauryl CoA, owing to the presence of a long hydrophobic tunnel that can accommodate C-12 chain of the substrate (10,30,32,38). Besides these three classes of acyl-CoA substrates (linear, branched and long chain linear acyl-CoAs), KASIII activity has not been explored with other types of substrates, such as hydroxylated, unsaturated or aromatic acyl-CoA thioesters. However, presence of ω -alicyclic (41,44,45,61), ω -halogenated (62) and ω -phenyl fatty acids in certain bacteria and plants suggests that KASIII enzymes from such organisms may have capabilities to utilize substrates other than linear or branched chain acyl-CoAs. In the current study, our data clearly validated the proposition that KASIII enzymes from certain organisms, especially from bacteria with high branched chain fatty acid content, can utilize unusual primers, such as hydroxylated acyl-CoAs (3-hydroxy butyryl CoA), aromatic acyl-CoAs (benzoyl CoA, phenylacetyl-CoA) and even unsaturated acyl-CoAs (crotonyl CoA).

New *in vivo* and *in vitro* screens for KASIII characterization

Traditionally, low through-put filter disc assays or conformationally sensitive urea gels have been used to assay KASIII activity using ^{14}C -labeled starter units (either ^{14}C -acyl-CoA or ^{14}C -malonyl ACP) (9). Herein we report both an *in vivo* and an *in vitro* method for rapidly screening KASIII substrate specificity with multiple substrates. The *in-vivo* screen is based on lethal *fabH* deletion mutant of *B. subtilis* bacteria which otherwise in its wild-type form makes 95% branched chain fatty acids (57,61). This lethal mutant can be rescued by either supplying branched chain fatty acids in the growth medium, or by introducing an exogenous KASIII gene into it which is capable of utilizing branched chain acyl-CoA primers (Jin et al, unpublished data). It is possible to detect KASIII functionality with other

acyl-CoA substrates by providing appropriate precursor carboxylic acids to the *B. subtilis* fabH deletion mutant. Our work, therefore, has established a novel *in-vivo* screening system to identify KASIII enzymes which have high specificity for branched chain primers and may also possess interesting substrate specificities.

The fluorescence-based thermal shift assay to survey the range of KASIII substrate specificity serves as a sensitive, rapid and reliable *in-vitro* screen, and complements the *B. subtilis* *in-vivo* screen to identify KASIII enzymes with novel substrate preferences. This technique was originally developed for high-throughput drug discovery (63,64) and recently, it has gained popularity as a quick method to determine optimal conditions for enhancing protein stability and crystallization (55,56,59). The current study illustrates the use of the thermofluor-based approach to biophysically characterize KASIII enzymes in a rapid and reliable manner.

Preference for unusual substrates by select KASIII enzymes

The current study establishes that several KASIII enzymes are capable of utilizing atypical substrates that do not naturally occur in many bacteria or plants (e.g. aromatic, unsaturated and hydroxylated acyl-CoAs). Our results demonstrate that KASIIIs derived from bacteria that produced branched chain or alicyclic fatty acids (such as *B. subtilis* KASIIb, *C. gingivalis* KASIIa, *A. acidocaldarius* KASIII and *T. aquaticus* KASIII) have a propensity for accepting atypical substrates as compared to the KASIII from *E. coli*, which does not synthesize branched chain fatty acids.

Overall, specific structural features of these KASIIIs that enable binding and catalysis of atypical substrates have not yet been determined. A possible explanation could be presence of a bigger hydrophobic tunnel that can accommodate the bigger, bulkier substrates.

A common feature of these four enzymes that could accept atypical substrates was the presence of the hinging Phe suggesting that they belong to group 3 KASIII enzymes which exhibit broad substrate specificities with higher preference for branched chain acyl-CoAs (chapter 3). The four enzymes, aaKASIII, taKASIII and cgKASIIIa showed less binding with short straight chain substrates such as acetyl- and propionyl-CoA and more interaction with branched chain acyl-CoAs. This is consistent with presence of high content of branched chain fatty acids in their host bacteria. Additionally, loops L3 and L9 of aaKASIII and taKASIII resemble those of group 3 KASIII enzymes. For example, similar to other group 3 KASIIIs that exhibit broad substrate specificities, aaKASIII and taKASIII have conserved Pro and Asp which are absent in loop L3 of group 1a enzymes that exhibit high specificity for straight chain primers (chapter 3 of this dissertation). Similarly, in loop L9 of aaKASIII and taKASIII, there is bulky Tyr/Phe in place of Thr190 of *E. coli* KASIII. Similarities in the L3, L9 loop regions of these enzymes with other group 3 KASIII enzymes suggest that other enzymes from group 3 may also be capable of utilizing atypical substrates.

A. acidocaldarius KASIII was the only KASIII that showed binding with benzoyl CoA, whereas both *B. subtilis* KASIIIb and *A. acidocaldarius* KASIII bound to phenylacetyl-CoA. Specific structural differences that allow *B. subtilis* KASIIIb to interact with phenylacetyl-CoA and not with benzoyl CoA can be understood in future by comparing its crystal structure with *A. acidocaldarius* KASIII's structure. Additionally, *A. acidocaldarius* KASIII bound with hexanoyl-CoA, whereas *C. gingivalis* KASIIIa, *B. subtilis* KASIIIb and *T. aquaticus* KASIII could not. This could be explained on the substitution of Val91 in *C. gingivalis* KASIIIa, Phe88 in *B. subtilis* KASIII and Phe86 in *T. aquaticus* KASIII with Leu88 in *A. acidocaldarius* KASIII that may be allowing longer acyl-CoA chains to fit in the

substrate binding pocket. The corresponding Phe87 in *E. coli* KASIII that extends from one subunit of the KASIII dimer into the acyl-CoA binding pocket of the other subunit is thought to limit substrate specificity to short chain acyl-CoAs (3,4,8). Previous studies have revealed that ability of *M. tuberculosis* KASIII and *M. luteus* KASIII to accept longer acyl-CoA substrates emanates from presence of a smaller Thr87 instead of Phe87 found in *E. coli* KASIII (10,17,30).

New possibilities for bio-based synthesis of novel functionalized fatty acids.

Typically bacteria and plants synthesize either straight or branched chain fatty acids initiated by the presence of straight or branched chain acyl-CoA primers. With the identification of KASIII enzymes that can process unusual CoA starter units, it is possible, as we have demonstrated, to synthesize novel ω -functionalized fatty acids that have different functional groups (such as hydroxyl, aromatic) at their terminal ends. Isolation and structural-functional characterization of KASIII genes from organisms like *Streptomyces cattleya*, *Dichapetalum toxicarium* which synthesize ω -fluorinated fatty acids, will help in further understanding of how KASIIIs are tailored to process halogenated substrates. Potential applications of such ω -functionalized fatty acids are in the bio-based chemical industry for production of various polymeric compounds (such as polyesters, polyamines, polystyrenes) and even lubricants and surfactants manufacturing industry. Such bio-based chemical intermediates can eventually replace the non-renewable petroleum based carbon feedstock in the current chemical industry and help in transitioning towards a bio-based economy.

Acknowledgements

This work was supported by the NSF ERC grant for Center for Biorenewable Chemicals (CBiRC), and was also supported in part by the CBiRC SLC sponsored grant. We would like to thank undergraduate students, Ruben Garcia and Morgan Becker, for their help in protein expression, purification and enzyme assays. We would also like to acknowledge the W.M. Keck Metabolomics lab where enzyme assays were conducted.

Table 1. Primer pairs for DNA manipulation in *B. subtilis*

Targeted amplification fragment	Vector	Primer name	Primer sequence
897-bp upstream fragment of the <i>bfabHA</i> ORF	pMU4A	AUf- <i>Pac</i> I	TTAATTAATATTAACCATCA CGGTGCAA
		AUr- <i>Sal</i> I	GTCGACGAATGTAACGTCCA ACACCA
799-bp downstream fragment of the <i>bfabHA</i> ORF	pMU4A	ADf- <i>Sal</i> I	GTCGACTGGAAGCCGGTAA AATCAA
		ADr- <i>Pst</i> I	CTGCAGGCCGACAATTTCTC CGTAAA
836-bp upstream fragment of the <i>bfabHB</i> ORF	pMU4B	BUf- <i>Pst</i> I	CTGCAGATATAAAACCGCCG GGACAT
		BUr- <i>Sal</i> I	GTCGACGCATAGGTGCCGAT AGCTGTA
802-bp downstream fragment of the <i>bfabHB</i> ORF	pMU4B	BDf- <i>Sal</i> I	GTCGACTCAAATCGTTTTGC TTTTCG
		BDr- <i>Pac</i> I	TTAATTAACCAAACAGGAGA TATCGATGC
836-bp upstream fragment of the <i>bfabHB</i> ORF	pUCB-erm	BUf2- <i>Eco</i> RI	GAATTCATATAAAACCGCCG GGACAT
		BUr2- <i>Sal</i> I	GTCGACGCATAGGTGCCGAT AGCTGTAA
738-bp downstream fragment of the <i>bfabHB</i> ORF	pUCB-erm	BDf2- <i>Sal</i> I	GTCGACTCAAATCGTTTTGC TTTTCG
		BDr2- <i>Hin</i> dIII	AAGCTTCCAAAGATGATGCC ATTCA
<i>erm</i> gene fragment	pUCB-erm	ermf	GTCGACCAAATTTACAAAAG CGACTCA
		ermr	GTCGACGAGGCCCTTTCGTC TTCAA
Verification of <i>bfabHA</i> deletion	Not applicable	lofAf	GCATACGCCTCCTTTCCATA
		lofAr	TTTGCCGGATATTCTTCAGC
Verification of <i>bfabHB</i> deletion	Not applicable	lofBf	CAATGTTAAGCCGGAAGGA A
		lofBr	AGCAGCCGTAAATGCCATA C

Table 2. Fatty acid production in *B. subtilis fabH* deletion mutants expressing exogenous KASIII enzymes

Source of KASIII	Fatty acid content (mole %)					
	Anteiso-fatty acids	Even-number iso-fatty acids	Odd-number iso-fatty acids	Even-number normal fatty acids	Odd-number normal fatty acids	Total branched chain fatty acids
<i>C. gingivalis</i> KASIIIa	27.24	1.28	58.94	12.46	0.08	87.46
<i>T. aquaticus</i> KASIII	52.1	1.09	26.89	19.84	0.12	80.08
<i>A. acidocaldarius</i> KASIII	48.13	3.64	30.44	17.73	0.05	82.21
<i>B. subtilis</i> KASIIIb*	54	13	28	5	0	95
<i>E. coli</i> KASIII*	0	0	0	95	5	0

* - Ratledge & Wilkinson, Microbial Lipids, Volume 1

Table 3. Melting temperatures (T_m) of the selected KASIII enzymes without ligand. These temperatures were used as the baseline temperatures against which the effect of addition of different substrates was measured.

Organism	Enzyme	Melting temperature (°C)
<i>C. gingivalis</i>	KASIIIa	68.3 ± 0.17
	KASIIIb	67.8 ± 0.03
	KASIIIc	56.8 ± 0.62
<i>A. acidocaldarius</i>	KASIII	73.2 ± 0.05
<i>T. aquaticus</i>	KASIII	84.1 ± 0.11
<i>B. subtilis</i>	KASIIIb	48.7 ± 0.24
<i>E. coli</i>	KASIII	55.0 ± 0.18

Bacterial source of KASIII	Branched chain fatty acids (%)	Active Site Residues			CoA binding residues			ACP binding residues					Residues affecting active site cleft shape		
		112	244	274	28	32	151	36	213	214	249	253	304	215	220
<i>Escherichia coli</i>	0	C	H	N	T	W	R	R	F	K	R	A	F	V	L
<i>Capnocytophaga gingivalis</i> KASIIIa	84	C	H	N	T	W	R	R	F	K	R	A	F	Y	M
<i>Capnocytophaga gingivalis</i> KASIIIb	84	C	H	N	T	W	R	R	F	K	R	F	F	N	F
<i>Capnocytophaga gingivalis</i> KASIIIc	84	C	H	N	W	K	K	K	F	N	H	K	F	F	T
<i>Thermus aquaticus</i>	95	C	H	N	T	W	R	R	F	K	R	A	F	F	M
<i>Alicyclobacillus acidocaldarius</i>	36	C	H	N	T	W	R	R	F	K	R	A	F	L	M
<i>B. subtilis</i> KASIIIb	95	C	H	N	T	W	R	R	Y	K	R	S	F	W	V

Figure 1. Alignment of conserved residues of the five putative KASIII enzymes with KASIII enzymes from both *E. coli* (narrow substrate specificity KASIII) and *B. subtilis* (broad substrate specificity KASIII). Fourteen conserved residues that define KASIII catalytic site, CoA and ACP binding sites and active site cleft shape are grouped into four categories respectively. These residues are compared for experimentally confirmed KASIII (E. coli narrow substrate specificity KASIII and B. subtilis broad substrate specificity KASIIIb) and five putative KASIII enzymes. Branched chain fatty acid content of each bacterium is listed. Coloring scheme for residues: Black – strictly conserved residues, Gray – homologous substitutions for the conserved residues, Blue – residues similar to those in *E. coli* KASIII but not in *B. subtilis* KASIIIb, Green – residues similar to those in *B. subtilis* KASIIIb but not in *E. coli* KASIII, White – non-conserved residues. Numbering scheme corresponds to the *E. coli* KASIII residues.

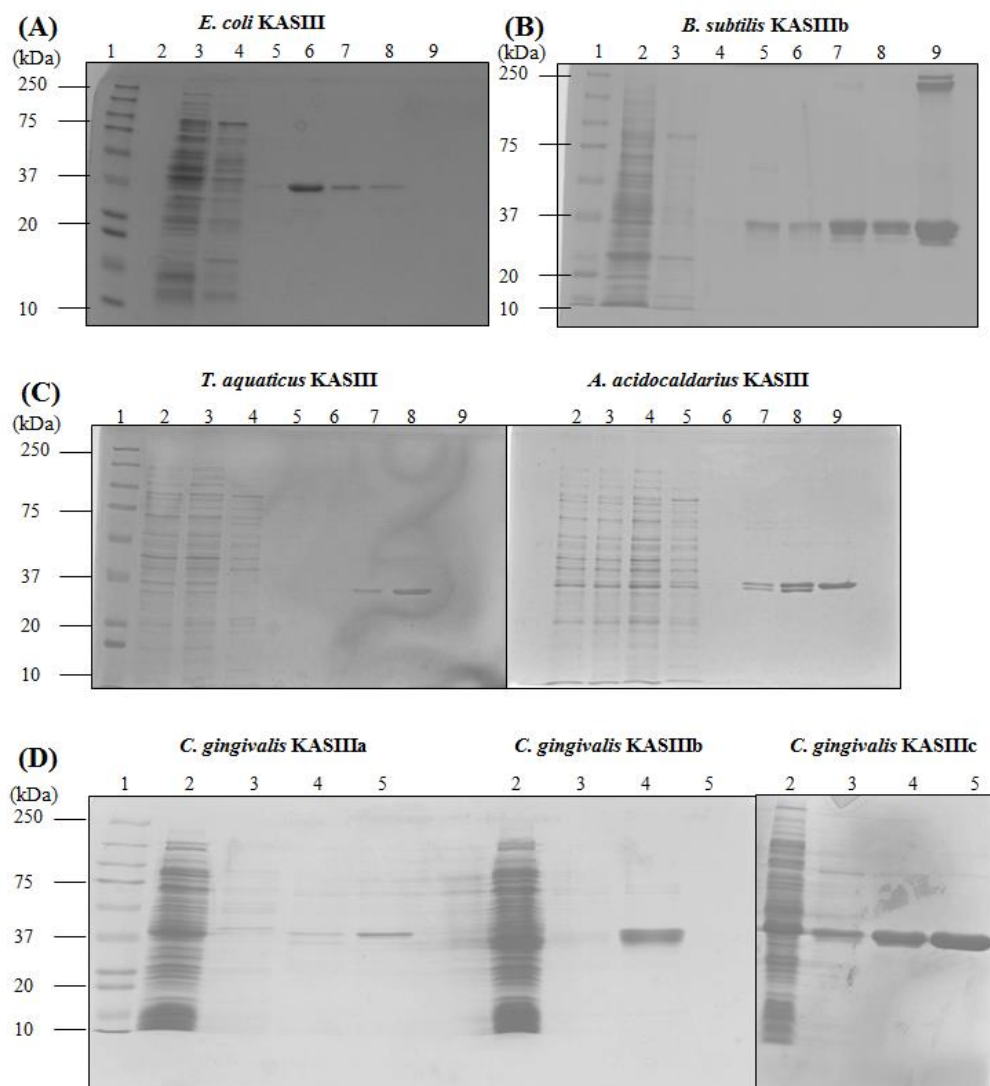


Figure 2. Expression and purification of the candidate KASIII proteins. Each of the five putative His-tagged KASIII proteins were purified by Ni-NTA affinity chromatography and assessed for homogeneity and purity using Coomassie-stained SDS PAGE gels. **(A)** Purified *E. coli* KASIII (1: Protein ladder, 2: Empty lane, 3: Crude extract, 4: Flow through, 5: First wash with 20 mM imidazole, 6: Wash with 40 mM imidazole buffer, 7: Elution with 80 mM imidazole, 8: Elution with 250 mM imidazole, 9: Empty lane). **(B)** Purified *B. subtilis* KASIIIb (1: Protein ladder, 2: Crude extract, 3: Flow through, 4: First wash with 20 mM imidazole, 5: Second wash with 20 mM imidazole, 6: Wash with 40 mM imidazole, 7: Elution with 80 mM imidazole, 8: Second elution with 80 mM imidazole, 9: Elution with 250 mM imidazole). **(C)** Purified *T. aquaticus* and *A. acidocaldarius* KASIIIs (1: Protein ladder, 2: Crude extract, 3: Flow through, 4: First wash with 20 mM imidazole, 5: Second wash with 20 mM imidazole, 6: Wash with 40 mM imidazole, 7: Elution with 80 mM imidazole, 8: Second elution with 80 mM imidazole, 9: Elution with 250 mM imidazole). **(D)** Purified *C. gingivalis* KASIIIa, KASIIIb and KASIIIc (1: Protein ladder, 2: Crude extract, 3: First wash, 4: Elution with 80 mM imidazole, 5: Elution with 250 mM imidazole).

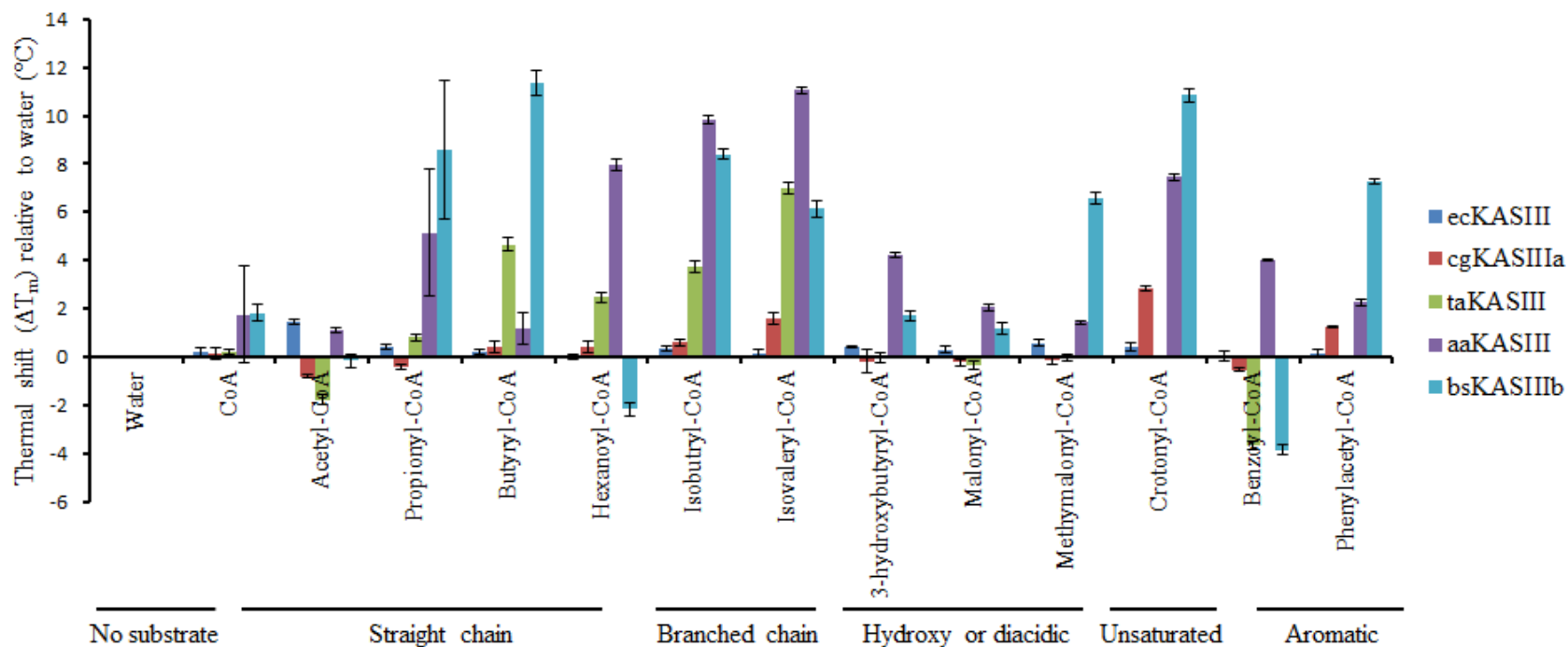


Figure 3. Binding of straight, branched, hydroxy, acidic, unsaturated and aromatic acyl-CoA substrates with the selected KASIIIs. Change in the melting temperatures (T_m) of the selected KASIIIs with respect to the baseline T_m upon addition of different acyl-CoA substrates is measured using fluorescence-based thermal shift assays. ecKASIII: *E. coli* KASIII, cgKASIIIa: *C. gingivalis* KASIIIa, taKASIII: *T. aquaticus* KASIII, aaKASIII: *A. acidocaldarius* KASIII, bsKASIIIb: *B. subtilis* KASIIIb. Each data point represents average of four technical replicates.

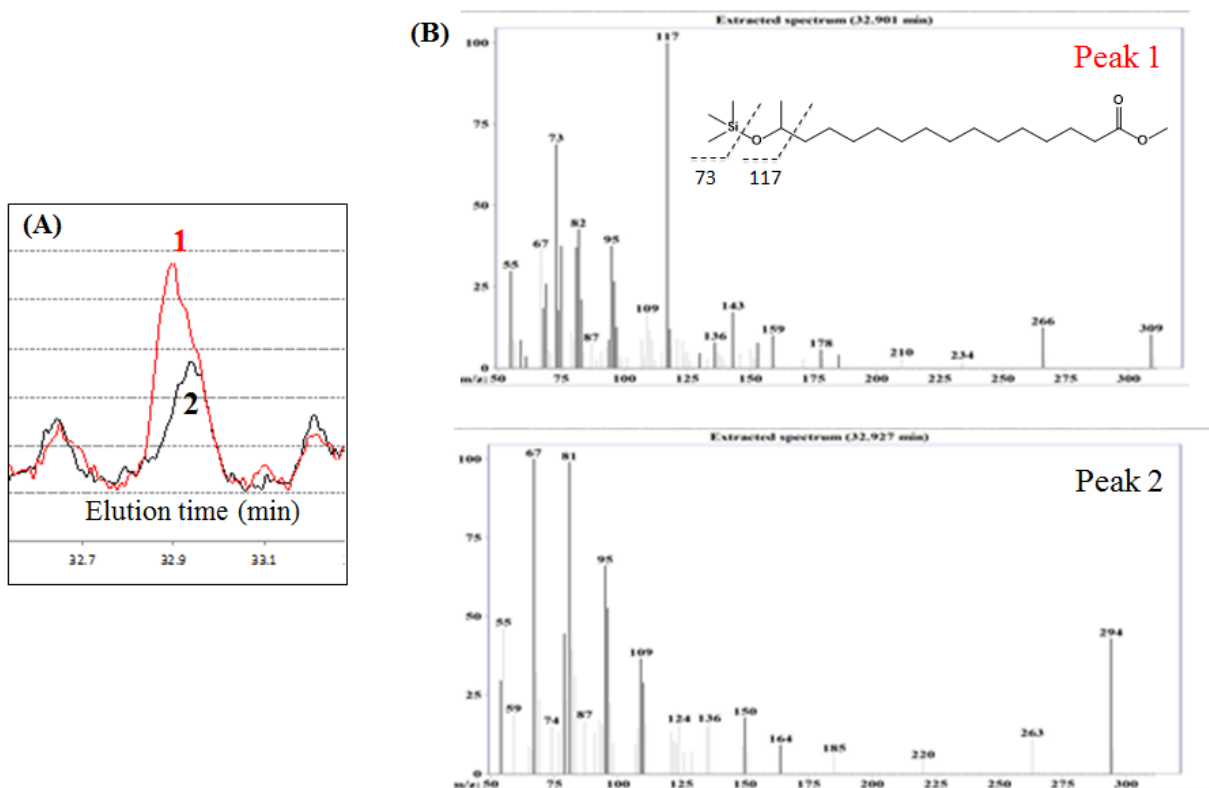


Figure 4. Production of ω -1-hydroxy-fatty acids by the expression of the *A. acidocaldarius* KASIII in *R. rubrum*. (A) GC-profile of the fatty acid products by the recombinant *R. rubrum* strain lacking PhaC activity, but expressing the *A. acidocaldarius* KASIII (red-line); the black-line is the parallel analysis of the profile from the control *R. rubrum* strain, lacking the *A. acidocaldarius* KASIII. (B). Mass-spectra of Peak 1 and Peak 2. Consistent with the analytical chemistry used to prepare these samples, Peak 1 is identified as the silylated-form of the methyl-ester of 15-hydroxypalmitate, while the absence of this metabolite in the control (Peak 2) established that the *A. acidocaldarius* KASIII is responsible for the production of (ω -1)-hydroxy-fatty acids.

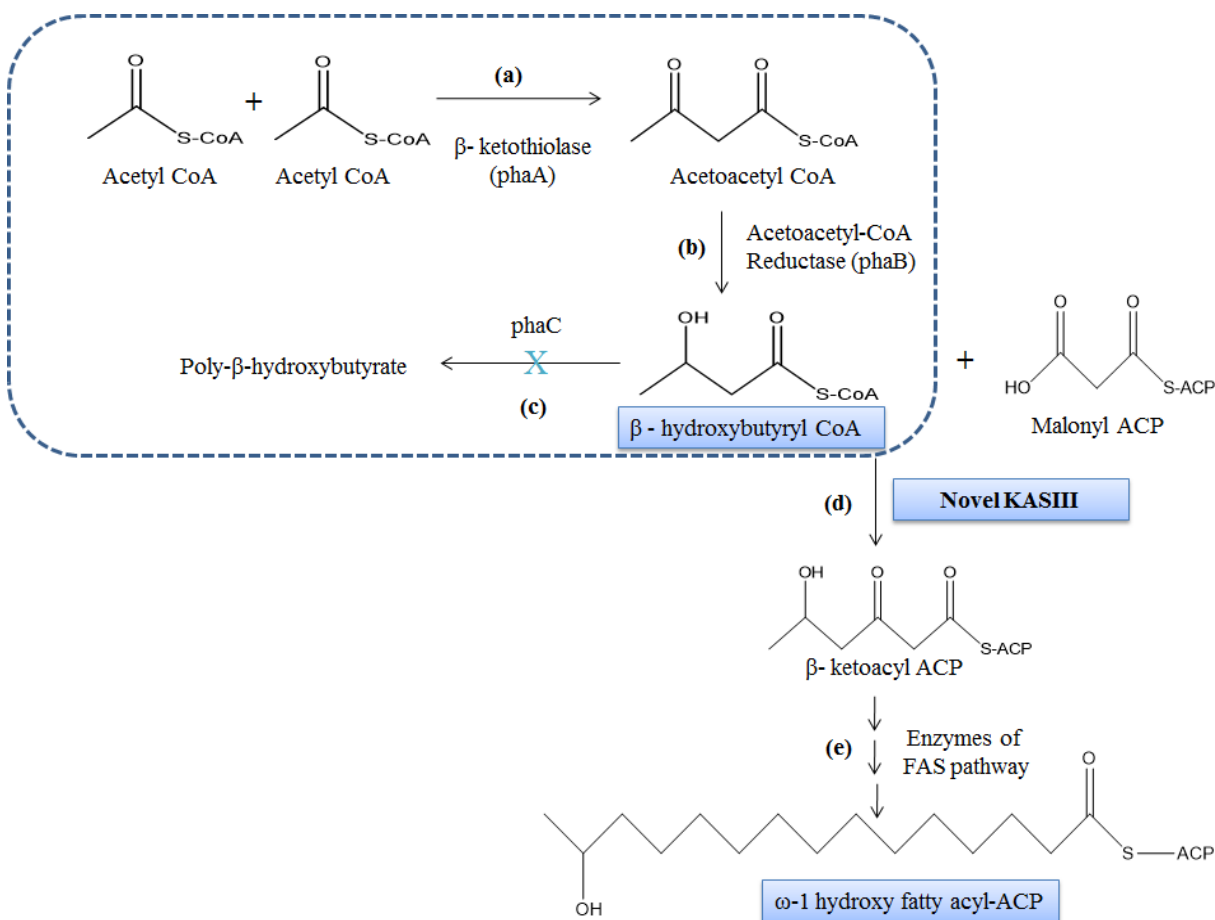


Figure 5. Engineered pathway for production of (ω -1)-hydroxy-fatty acids in *R. rubrum* strain. Poly- β -hydroxybutyrate (PBHB) is synthesized in *R. rubrum* via condensation of two acetyl-CoA molecules by β -ketothiolase (phaA) to form acetoacetyl-CoA (a), followed by reduction of acetoacetyl-CoA to β -hydroxybutyryl CoA by acetoacetyl-CoA reductase (phaB) (b), which is then polymerized by phaC to form PBHB (c). In a *R. rubrum* strain that lacks phaC, *A. acidocaldarius* KASIII was introduced that can condense β -hydroxybutyryl CoA with malonyl-ACP to form a hydroxylated β -ketoacyl-ACP (d), which can be then elongated by enzymes of fatty acid synthesis pathway (e) to synthesize ω -1-hydroxy-fatty acid.

References

1. Lu, Y. J., Zhang, Y. M., and Rock, C. O. (2004) Product diversity and regulation of type II fatty acid synthases. *Biochem Cell Biol* 82, 145-155
2. Heath, R. J., and Rock, C. O. (2002) The Claisen condensation in biology. *Nat Prod Rep* 19, 581-596
3. Qiu, X., Janson, C. A., Konstantinidis, A. K., Nwagwu, S., Silverman, C., Smith, W. W., Khandekar, S., Lonsdale, J., and Abdel-Meguid, S. S. (1999) Crystal structure of beta-ketoacyl-acyl carrier protein synthase III. A key condensing enzyme in bacterial fatty acid biosynthesis. *J Biol Chem* 274, 36465-36471
4. Qiu, X., Janson, C. A., Smith, W. W., Head, M., Lonsdale, J., and Konstantinidis, A. K. (2001) Refined structures of beta-ketoacyl-acyl carrier protein synthase III. *J Mol Biol* 307, 341-356
5. Qiu, X., Choudhry, A. E., Janson, C. A., Grooms, M., Daines, R. A., Lonsdale, J. T., and Khandekar, S. S. (2005) Crystal structure and substrate specificity of the beta-ketoacyl-acyl carrier protein synthase III (FabH) from *Staphylococcus aureus*. *Protein Sci* 14, 2087-2094
6. Tsay, J. T., Oh, W., Larson, T. J., Jackowski, S., and Rock, C. O. (1992) Isolation and characterization of the beta-ketoacyl-acyl carrier protein synthase III gene (fabH) from *Escherichia coli* K-12. *J Biol Chem* 267, 6807-6814
7. Han, L., Lobo, S., and Reynolds, K. A. (1998) Characterization of beta-ketoacyl-acyl carrier protein synthase III from *Streptomyces glaucescens* and its role in initiation of fatty acid biosynthesis. *J Bacteriol* 180, 4481-4486
8. Davies, C., Heath, R. J., White, S. W., and Rock, C. O. (2000) The 1.8 Å crystal structure and active-site architecture of beta-ketoacyl-acyl carrier protein synthase III (FabH) from *Escherichia coli*. *Structure* 8, 185-195
9. Choi, K. H., Heath, R. J., and Rock, C. O. (2000) beta-ketoacyl-acyl carrier protein synthase III (FabH) is a determining factor in branched-chain fatty acid biosynthesis. *J Bacteriol* 182, 365-370
10. Choi, K. H., Kremer, L., Besra, G. S., and Rock, C. O. (2000) Identification and substrate specificity of beta -ketoacyl (acyl carrier protein) synthase III (mtFabH) from *Mycobacterium tuberculosis*. *J Biol Chem* 275, 28201-28207
11. Khandekar, S. S., Konstantinidis, A. K., Silverman, C., Janson, C. A., McNulty, D. E., Nwagwu, S., Van Aller, G. S., Doyle, M. L., Kane, J. F., Qiu, X., and Lonsdale, J. (2000) Expression, purification, and crystallization of the *Escherichia coli* selenomethionyl beta-ketoacyl-acyl carrier protein synthase III. *Biochem Biophys Res Commun* 270, 100-107

12. Khandekar, S. S., Gentry, D. R., Van Aller, G. S., Warren, P., Xiang, H., Silverman, C., Doyle, M. L., Chambers, P. A., Konstantinidis, A. K., Brandt, M., Daines, R. A., and Lonsdale, J. T. (2001) Identification, substrate specificity, and inhibition of the *Streptococcus pneumoniae* beta-ketoacyl-acyl carrier protein synthase III (FabH). *J Biol Chem* 276, 30024-30030
13. Huynh, K. H., Natarajan, S., Song, N. H., Ngo, P. T., Ahn, Y. J., Kim, J. G., Lee, B. M., Eo, Y. D., and Kang, L. W. (2009) Cloning, expression, crystallization and preliminary X-ray crystallographic analysis of beta-ketoacyl-ACP synthase III (FabH) from *Xanthomonas oryzae* pv. *oryzae*. *Acta Crystallogr Sect F Struct Biol Cryst Commun* 65, 460-462
14. Revill, W. P., Bibb, M. J., Scheu, A. K., Kieser, H. J., and Hopwood, D. A. (2001) Beta-ketoacyl acyl carrier protein synthase III (FabH) is essential for fatty acid biosynthesis in *Streptomyces coelicolor* A3(2). *J Bacteriol* 183, 3526-3530
15. Singh, A. K., Zhang, Y. M., Zhu, K., Subramanian, C., Li, Z., Jayaswal, R. K., Gatto, C., Rock, C. O., and Wilkinson, B. J. (2009) FabH selectivity for anteiso branched-chain fatty acid precursors in low-temperature adaptation in *Listeria monocytogenes*. *FEMS Microbiol Lett* 301, 188-192
16. Wen, L., Chmielowski, J. N., Bohn, K. C., Huang, J. K., Timsina, Y. N., Kodali, P., and Pathak, A. K. (2009) Functional expression of *Francisella tularensis* FabH and FabI, potential antibacterial targets. *Protein Expr Purif* 65, 83-91
17. Pereira, J. H., Goh, E. B., Keasling, J. D., Beller, H. R., and Adams, P. D. (2012) Structure of FabH and factors affecting the distribution of branched fatty acids in *Micrococcus luteus*. *Acta Crystallogr D Biol Crystallogr* 68, 1320-1328
18. Clough, R. C., Matthis, A. L., Barnum, S. R., and Jaworski, J. G. (1992) Purification and characterization of 3-ketoacyl-acyl carrier protein synthase III from spinach. A condensing enzyme utilizing acetyl-coenzyme A to initiate fatty acid synthesis. *J Biol Chem* 267, 20992-20998
19. Jaworski, J. G., Clough, R. C., and Barnum, S. R. (1989) A Cerulenin Insensitive Short Chain 3-Ketoacyl-Acyl Carrier Protein Synthase in *Spinacia oleracea* Leaves. *Plant Physiol* 90, 41-44
20. Tai, H., Post-Beittenmiller, D., and Jaworski, J. G. (1994) Cloning of a cDNA encoding 3-ketoacyl-acyl carrier protein synthase III from *Arabidopsis*. *Plant Physiol* 106, 801-802
21. Dehesh, K., Tai, H., Edwards, P., Byrne, J., and Jaworski, J. G. (2001) Overexpression of 3-ketoacyl-acyl-carrier protein synthase IIIs in plants reduces the rate of lipid synthesis. *Plant Physiol* 125, 1103-1114

22. Abbadi, A., Brummel, M., and Spener, F. (2000) Knockout of the regulatory site of 3-ketoacyl-ACP synthase III enhances short- and medium-chain acyl-ACP synthesis. *Plant J* 24, 1-9
23. Abbadi, A., Brummel, M., Schütt, B. S., Slabaugh, M. B., Schuch, R., and Spener, F. (2000) Reaction mechanism of recombinant 3-oxoacyl-(acyl-carrier-protein) synthase III from *Cuphea wrightii* embryo, a fatty acid synthase type II condensing enzyme. *Biochem J* 345 Pt 1, 153-160
24. Li, J., Li, M. R., Wu, P. Z., Tian, C. E., Jiang, H. W., and Wu, G. J. (2008) Molecular cloning and expression analysis of a gene encoding a putative beta-ketoacyl-acyl carrier protein (ACP) synthase III (KAS III) from *Jatropha curcas*. *Tree Physiol* 28, 921-927
25. González-Mellado, D., von Wettstein-Knowles, P., Garcés, R., and Martínez-Force, E. (2010) The role of beta-ketoacyl-acyl carrier protein synthase III in the condensation steps of fatty acid biosynthesis in sunflower. *Planta* 231, 1277-1289
26. Prigge, S. T., He, X., Gerena, L., Waters, N. C., and Reynolds, K. A. (2003) The initiating steps of a type II fatty acid synthase in *Plasmodium falciparum* are catalyzed by pfACP, pfMCAT, and pfKASIII. *Biochemistry* 42, 1160-1169
27. Waters, N. C., Kopydlowski, K. M., Guszczynski, T., Wei, L., Sellers, P., Ferlan, J. T., Lee, P. J., Li, Z., Woodard, C. L., Shallom, S., Gardner, M. J., and Prigge, S. T. (2002) Functional characterization of the acyl carrier protein (PfACP) and beta-ketoacyl ACP synthase III (PfKASIII) from *Plasmodium falciparum*. *Mol Biochem Parasitol* 123, 85-94
28. Gajiwala, K. S., Margosiak, S., Lu, J., Cortez, J., Su, Y., Nie, Z., and Appelt, K. (2009) Crystal structures of bacterial FabH suggest a molecular basis for the substrate specificity of the enzyme. *FEBS Lett* 583, 2939-2946
29. Scarsdale, J. N., Kazanina, G., He, X., Reynolds, K. A., and Wright, H. T. (2001) Crystal structure of the *Mycobacterium tuberculosis* beta-ketoacyl-acyl carrier protein synthase III. *J Biol Chem* 276, 20516-20522
30. Musayev, F., Sachdeva, S., Scarsdale, J. N., Reynolds, K. A., and Wright, H. T. (2005) Crystal structure of a substrate complex of *Mycobacterium tuberculosis* beta-ketoacyl-acyl carrier protein synthase III (FabH) with lauroyl-coenzyme A. *J Mol Biol* 346, 1313-1321
31. Sachdeva, S., Musayev, F. N., Alhamadsheh, M. M., Scarsdale, J. N., Wright, H. T., and Reynolds, K. A. (2008) Separate entrance and exit portals for ligand traffic in *Mycobacterium tuberculosis* FabH. *Chem Biol* 15, 402-412
32. Brown, A. K., Sridharan, S., Kremer, L., Lindenberg, S., Dover, L. G., Sacchettini, J. C., and Besra, G. S. (2005) Probing the mechanism of the *Mycobacterium*

- tuberculosis beta-ketoacyl-acyl carrier protein synthase III mtFabH: factors influencing catalysis and substrate specificity. *J Biol Chem* 280, 32539-32547
33. Daines, R. A., Pendrak, I., Sham, K., Van Aller, G. S., Konstantinidis, A. K., Lonsdale, J. T., Janson, C. A., Qiu, X., Brandt, M., Khandekar, S. S., Silverman, C., and Head, M. S. (2003) First X-ray cocrystal structure of a bacterial FabH condensing enzyme and a small molecule inhibitor achieved using rational design and homology modeling. *J Med Chem* 46, 5-8
 34. Heath, R. J., and Rock, C. O. (1996) Inhibition of beta-ketoacyl-acyl carrier protein synthase III (FabH) by acyl-acyl carrier protein in *Escherichia coli*. *J Biol Chem* 271, 10996-11000
 35. Jackowski, S., Murphy, C. M., Cronan, J. E., and Rock, C. O. (1989) Acetoacetyl-acyl carrier protein synthase. A target for the antibiotic thiolactomycin. *J Biol Chem* 264, 7624-7629
 36. Jackowski, S., and Rock, C. O. (1987) Acetoacetyl-acyl carrier protein synthase, a potential regulator of fatty acid biosynthesis in bacteria. *J Biol Chem* 262, 7927-7931
 37. Smirnova, N., and Reynolds, K. A. (2001) Branched-chain fatty acid biosynthesis in *Escherichia coli*. *J Ind Microbiol Biotechnol* 27, 246-251
 38. Sachdeva, S., and Reynolds, K. A. (2008) *Mycobacterium tuberculosis* beta-ketoacyl acyl carrier protein synthase III (mtFabH) assay: principles and method. *Methods Mol Med* 142, 205-213
 39. PETERS, R. A., HALL, R. J., WARD, P. F., and SHEPPARD, N. (1960) The chemical nature of the toxic compounds containing fluorine in the seeds of *Dichapetalum toxicarium*. *Biochem J* 77, 17-23
 40. Meija, J., and Soukup, V. G. (2004) Phenyl-terminated fatty acids in seeds of various aroids. *Phytochemistry* 65, 2229-2237
 41. De Rosa, M., Gambacorta, A., Minale, L., and Bu'lock, J. D. (1972) The formation of -cyclohexyl-fatty acids from shikimate in an acidophilic thermophilic bacillus. A new biosynthetic pathway. *Biochem J* 128, 751-754
 42. De Rosa, M., Gambacorta, A., and Bu'lock, J. D. (1974) Effects of pH and temperature on the fatty acid composition of *Bacillus acidocaldarius*. *J Bacteriol* 117, 212-214
 43. Darland, G., and Brock, T. D. (1971) *Bacillus acidocaldarius* sp. nov., an acidophilic thermophilic spore-forming bacterium., *J. Gen. Microbiol.*
 44. Siristova, L., Luhovy, R., Sigler, K., and Rezanka, T. (2011) Biosynthesis of ω -alicyclic fatty acids induced by cyclic precursors and change of membrane fluidity in

- thermophilic bacteria *Geobacillus stearothermophilus* and *Meiothermus ruber*. *Extremophiles* 15, 423-429
45. Dreher, R., Poralla, K., and König, W. A. (1976) Synthesis of omega-alicyclic fatty acids from cyclic precursors in *Bacillus subtilis*. *J Bacteriol* 127, 1136-1140
 46. Jin, H., and Nikolau, B. J. (2012) Role of genetic redundancy in polyhydroxyalkanoate (PHA) polymerases in PHA biosynthesis in *Rhodospirillum rubrum*. *J Bacteriol* 194, 5522-5529
 47. Nikolau, B. J., Perera, M. A., Brachova, L., and Shanks, B. (2008) Platform biochemicals for a biorenewable chemical industry. *Plant J* 54, 536-545
 48. Vagner, V., Dervyn, E., and Ehrlich, S. D. (1998) A vector for systematic gene inactivation in *Bacillus subtilis*. *Microbiology* 144 (Pt 11), 3097-3104
 49. Ways, P., and Hanahan, D. J. (1964) Characterization and quantification of red cell lipids in normal man. *J Lipid Res* 5, 318-328
 50. Harvey, D. J. (1982) Picolinyl esters as derivatives for the structural determination of long chain branched and unsaturated fatty acids. *Biomed Mass Spectrom*
 51. Broekman, J. H., and Steenbakkers, J. F. (1973) Growth in high osmotic medium of an unsaturated fatty acid auxotroph of *Escherichia coli* K-12. *J Bacteriol* 116, 285-289
 52. Broekman, J. H., and Steenbakkers, J. F. (1974) Effect of the osmotic pressure of the growth medium on *fabB* mutants of *Escherichia coli*. *J Bacteriol* 117, 971-977
 53. Buser, H. R., Arn, H., and Guerlin, P. (1983) Determination of double bond position in mono-unsaturated acetates by mass spectrometry of dimethyl disulfides adducts. *Anal Chem* 55, 818-822
 54. National BioResource Project (NIG, Japan) : *E.coli*.
 55. Ericsson, U. B., Hallberg, B. M., Detitta, G. T., Dekker, N., and Nordlund, P. (2006) Thermofluor-based high-throughput stability optimization of proteins for structural studies. *Anal Biochem* 357, 289-298
 56. Niesen, F. H., Berglund, H., and Vedadi, M. (2007) The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nat Protoc* 2, 2212-2221
 57. Ratledge, C., and Wilkinson, S. G. (1988) *Microbial Lipids*, Academic Press, UK
 58. Rock, C. O., and Cronan, J. E. (1996) *Escherichia coli* as a model for the regulation of dissociable (type II) fatty acid biosynthesis. *Biochim Biophys Acta* 1302, 1-16

59. Vedadi, M., Niesen, F. H., Allali-Hassani, A., Fedorov, O. Y., Finerty, P. J., Wasney, G. A., Yeung, R., Arrowsmith, C., Ball, L. J., Berglund, H., Hui, R., Marsden, B. D., Nordlund, P., Sundstrom, M., Weigelt, J., and Edwards, A. M. (2006) Chemical screening methods to identify ligands that promote protein stability, protein crystallization, and structure determination. *Proc Natl Acad Sci U S A* 103, 15835-15840
60. Tsay, J. T., Oh, W., Larson, T. J., Jackowski, S., and Rock, C. O. (1992) Isolation and characterization of the beta-ketoacyl-acyl carrier protein synthase III gene (fabH) from *Escherichia coli* K-12. *J Biol Chem* 267, 6807-6814
61. Kaneda, T. (1991) Iso- and anteiso-fatty acids in bacteria: biosynthesis, function, and taxonomic significance. *Microbiol Rev* 55, 288-302
62. O' Hagan, D., and Harper, D. B. (1999) Fluorine-containing natural products. *Journal of fluorine chemistry*
63. Lo, M. C., Aulabaugh, A., Jin, G., Cowling, R., Bard, J., Malamas, M., and Ellestad, G. (2004) Evaluation of fluorescence-based thermal shift assays for hit identification in drug discovery. *Anal Biochem* 332, 153-159
64. Pantoliano, M. W., Petrella, E. C., Kwasnoski, J. D., Lobanov, V. S., Myslik, J., Graf, E., Carver, T., Asel, E., Springer, B. A., Lane, P., and Salemme, F. R. (2001) High-density miniaturized thermal shift assays as a general strategy for drug discovery. *J Biomol Screen* 6, 429-440

CHAPTER V

GENERAL CONCLUSIONS

This dissertation investigates the structure-function relationship of 3-Ketoacyl ACP Synthase III enzymes and explores their functional diversity. The two main goals of this dissertation were: 1) to understand the structure-function relationships of KASIII enzymes and identify key residues governing the KASIII substrate specificity, and 2) to explore the functional diversity of KASIII enzymes and identify KASIII enzymes with novel substrate specificities. Through this work, our fundamental understanding of the KASIII structure-function relationship has been enhanced and we have identified several structural determinants of KASIII substrate specificity. Moreover, we have identified several KASIIIs with novel substrate specificities. Taken together, this expanded understanding of KASIII will help us to expand the biosynthetic repertoire of the fatty acid synthesis pathway to make novel functionalized fatty acids that could have applications in the bio-based chemical industry (1). Presented herein are the main conclusions from this work.

The conserved Phe determines KASIII substrate specificity and is crucial for substrate binding and catalysis

Experimental results from chapter 2 demonstrated the importance of the conserved Phe and the two residues in its vicinity (corresponding to Phe304, Val 215 and Leu220 of *E. coli* KASIII that exhibits narrow substrate specificity for straight chain primers) in substrate binding and catalysis in addition to determining the substrate specificity of KASIII. *B. subtilis* KASIIIa and KASIIIb enzymes, which have broad substrate specificity for straight and branched chain primers, have the conserved Phe but the two neighboring residues that

are thought to affect its rotamer conformation are bulky Phe and Met in bsKASIIIa (Trp and Val in bsKASIIIb) respectively.

Previously it was believed that KASIII from Gram-negative bacteria possessed residues similar to the ones found in *E. coli* KASIII at these positions (i.e. Val215 and Leu220), and KASIII from Gram-positive bacteria had residues similar to the ones in bsKASIIIa (i.e. Phe and Met). Our phylogenetic analysis from chapter 3 and 4 clearly showed that irrespective of the Gram-character of the host bacterium, the KASIII had residues similar to *E. coli* KASIII if the host bacterium makes only straight chain fatty acids. In contrast, if the host bacterium synthesizes branched chain fatty acids or alicyclic fatty acids, then the residues neighboring the conserved Phe are similar to the ones occurring in *B. subtilis* KASIII enzymes.

Interestingly, we also identified a functional KASIII enzyme, *Myxococcus xanthus* KASIIIc, that does not have the ‘conserved’ Phe but still shows KASIII functionality and is more active with straight chain acyl-CoA substrates and has little activity with branched chain primers. Previously, only two KASIIIs have been structurally characterized, i.e. *A. aeolicus* KASIII (PDB code: 2EBD) and *X. oryzae* KASIII (2), that do not possess the conserved Phe. Functional characterization of these enzymes will be instrumental in determining if absence of the conserved Phe limits the KASIII activity to straight chain acyl-CoA substrates as observed in the case of *M. xanthus* KASIIIc.

In conclusion, we experimentally validated the importance of the conserved Phe and the two neighboring residues in determining KASIII substrate specificity and catalysis. We also identified five KASIII enzymes from Gram-negative and Gram-positive bacteria that have bulky residues close to conserved Phe, and therefore exhibit broad substrate specificity.

Future perspective – In addition to the residues studied in chapter 2, there may be other residues that determine KASIII substrate specificity. For example, Phe87 in *E. coli* KASIII at the base of the substrate binding pocket extends from one subunit of the KASIII dimer into the active site of the other subunit (3-5). In the KASIII from *M. tuberculosis*, which accepts longer acyl CoA substrates such as lauroyl CoA (6-8), this Phe residue is replaced by a smaller residue, Thr. Therefore, it is proposed that the larger Phe87 residue may be hindering entrance of longer acyl-CoA substrates into the *E. coli* KASIII substrate binding pocket. Likewise, Arg196 in *E. coli* KASIII, which also participates in dimer interface formation, may also have a role in limiting the specificity of ecKASIII to straight chain substrates, because this residue is mostly conserved in KASIIIs from bacteria that make straight chain fatty acids, such as *Haemophilus influenzae* and *Yersinia pestis*. The roles of these specific residues could be experimentally tested to attain better understanding of the KASIII structure-function relationship.

***E. coli* KASIII is capable of binding and processing branched chain primers, such as isobutyryl CoA**

E. coli KASIII is known to have high selectivity for acetyl-CoA (9-13), however there has been ambiguity regarding its capability to utilize branched chain substrates, such as isobutyryl-CoA. One study demonstrated that *E. coli* KASIII is incapable of utilizing isobutyryl-CoA (14). In contrast, another report demonstrated that *E. coli* KASIII, when supplied with appropriate branched-chain precursors, is capable of producing branched chain fatty acids, although at a rate ten-fold lower than for straight-chain fatty acids (15). Our experimental data from saturation transfer difference NMR experiments and enzymatic activity analysis (Chapter 2) supported the results of the second study (15), by showing that

E. coli KASIII is capable of binding isobutyryl-CoA. Via STD-NMR, epitopes were identified on isobutyryl-CoA that directly interact with the KASIII enzymes. These ligand binding epitopes are similar to the ones on the straight chain acyl-CoA ligand (acetyl-CoA), which suggests that the binding modes of straight and branched chain substrates are identical. The enzymatic activity data further illustrated that *E. coli* KASIII is capable of processing branched-chain substrates, but at a lower specific activity than straight chain substrates (e.g. acetyl CoA). Taken together, these results suggest that the formation of branched chain fatty acids in *E. coli* is not limited by the substrate specificity of KASIII alone, but also by the unavailability of branched-chain precursors *in vivo*.

Future perspective – Disordered form of *E. coli* KASIII has been observed in crystal structure suggesting large conformational changes in this enzyme and models have been suggested for *E. coli* KASIII conformational flexibility (16), so it is quite possible that binding of straight chain substrates induces conformational changes that branched chain substrates are incapable of inducing. Such molecular dynamics and conformational changes in KASIII structure have not been experimentally explored yet. Therefore, NMR based structures in solutions of these particular KASIII enzymes will be instrumental in advancing the existing structural knowledge of KASIII enzymes and could provide insights into the conformational changes induced in ecKASIII upon binding different ligands (i.e. straight vs. branched chain substrates).

KASIII can be categorized into four structure-function groups based on which their substrate specificity can be predicted

Collectively, in chapters 3 and 4, we identified and characterized six functional KASIII enzymes out of twelve previously uncharacterized KASIII enzymes from five diverse

sources. All these enzymes, specifically *C. gingivalis* KASIIIa, KASIIIc, *L. pneumophila* KASIIIa, *M. xanthus* KASIIIc, *A. acidocaldarius* KASIII and *T. aquaticus* KASIII, exhibited activity and binding towards both straight and branched chain substrates. In addition to the six KASIIIs functionally characterized here, functional data is available for nine other KASIIIs from different bacteria, i.e. *E. coli* (17-19), *B. subtilis* (11), *S. aureus* (20), *M. tuberculosis* (21), *M. luteus* (22), *L. monocytogenes* (23), *S. pneumoniae* (24), *S. glaucescens* (25), *F. tularensis* (26). Structural information, however, is available for KASIII from 12 diverse sources, which include *E. coli* (3-5,27), *M. tuberculosis* (6,7,28), *M. luteus* (22), *S. aureus* (27), *X. oryzae* (2), *B. pseudomallei* (PDB code: 3GWE), *A. aeolicus* (PDB code: 2EBD), *T. thermophilus* (PDB code: 1UB7), *H. influenza* (27), *E. faecalis* (27), *B. xenovorans* (PDB code: 4DFE) and *P. aeruginosa* (PDB code: 2X3E). There are only four KASIIIs from these two groups for which both structural and functional information is available, so there is still a huge gap between structural and functional features of KASIII enzymes.

We correlated the sequence and function information for all the characterized KASIIIs to date and categorized KASIII enzymes into four structure-function groups (chapter 3) based on primary sequence similarities and functional similarities. From this classification, two groups emerged that showed either high selectivity for straight chain primers (group 1a consisting of KASIII from *E. coli*, *Y. pestis*) or high selectivity for branched chain primers (group 3 consisting of *B. subtilis* KASIIIa, KASIIIb, *T. thermophilus* KASIII). Group 1b had KASIII that could accept both straight and branched chain primers but higher activity for the former substrates. Group 2 also had KASIII with similar substrate

preferences but it also contained some unique KASIIIs, such as *M. tuberculosis* and *M. luteus* KASIIIs that can bind to long chain acyl-CoA substrates (21,22).

Additionally, we also identified common structural features in the loops L3 and L9 of these four groups which could be responsible for conferring different substrate specificities to each of the groups. These two loops form the dimer interface of KASIII enzyme, and may therefore have an important role in inducing conformational changes in the KASIII structure upon binding of the appropriate substrates.

Future perspective –As mentioned before, there are only four KASIII enzymes for which both structural and functional data is available. To further enhance our understanding of KASIII enzymes, it is imperative that functional data is collected for the KASIIIs that have been crystallized, and likewise structural information should be gathered for the KASIIIs that have only been functionally characterized. Such efforts will tremendously help in answering the remaining questions about KASIII structure-function relationships. The most interesting candidates for structure-function studies would be *C. gingivalis* KASIIIa, *B. subtilis* KASIIIb, *A. acidocaldarius* KASIII that can bind to unusual substrates as demonstrated by our data from chapter 4.

Phylogenetic divergence of KASIIIs results in at least one functional homolog in each bacterial strain

In chapter 3 we characterized multiple KASIII homologs from three different bacterial sources and established that at least one KASIII homolog per strain was functional. In the case of KASIII homologs from *C. gingivalis*, cgKASIIIa and cgKASIIIc that belong to sub-families IV and X of the KASIII phylogenetic tree respectively, showed marked dissimilarities in their substrate preferences. cgKASIIIa is highly selective for branched-

chain acyl CoA substrates whereas cgKASIIIb has a stronger preference for straight chain primers, but can also utilize branched chain primers. In case of *L. pneumophila*, four KASIII paralogs belonged to different sub-families but only one of these exhibited KASIII functionality. Likewise, only mxKASIIIC from *M. xanthus* displayed KASIII activity and could process both straight and branched chain substrates, whereas mxKASIIIA and mxKASIIIB did not show KASIII functionality. In summary, each of these three bacteria had at least one functional KASIII that was active with branched chain substrates. This is consistent with the production of branched chain fatty acids in these bacteria, and sequence analysis of the functional KASIIIs revealed that these had features similar to those of other KASIIIs that are active with branched chain primers, such as *B. subtilis* KASIIIA and *S. aureus* KASIII.

Future perspective – The physiological significance of multiple functional KASIII homologs within a single bacterial strain is not yet understood. *Bacillus subtilis* is a very well characterized organism that has two KASIII homologs, bsKASIIIA and bsKASIIIB, that show differences in their substrate preferences (11). However, physiological significance of this duplication in function is still under investigation. Preliminary data from some studies suggest that bsKASIIIA is expressed constitutively whereas bsKASIIIB is expressed at lower temperatures, suggesting a role of the two KASIIIs in helping *Bacillus* adapt to temperature changes (Jin & Nikolau, unpublished data). Similar roles could be predicted for *C. gingivalis* KASIIIA and KASIIIC, i.e. cgKASIIIA that can utilize branched chain acyl CoAs might be expressed constitutively resulting in high branched chain fatty acid content of *C. gingivalis*, whereas cgKASIIIC could have non-constitutive expression. Gene knockout experiments would provide more insights into their physiological functions.

KASIII from specific organisms are capable of utilizing novel acyl CoA substrates

In chapter 4, four novel KASIII enzymes were shown to utilize at least one of several atypical acyl CoA substrates, including diacid (malonyl and methylmalonyl-CoA), hydroxylated (3-hydroxybutyryl-CoA), aromatic (benzoyl-CoA and phenylacetyl-CoA) and unsaturated (crotonyl-CoA) acyl-CoAs. These four enzymes belong to bacterial species that either produce high percentages of branched-chain fatty acids (*Bacillus subtilis*, *Capnocytophaga gingivalis*, *Thermus thermophilus* (29)) or make more than 50% alicyclic fatty acids (*Alicyclobacillus acidocaldarius*) (29). Typically, bacteria and plants synthesize either straight or branched chain fatty acids, initiated by the presence of straight or branched chain acyl CoA primers. With the identification of KASIII enzymes that can process unusual CoA starter units, it is possible to synthesize novel ω -functionalized fatty acids that have different functional groups (e.g. carboxylic acid or hydroxyl or aromatic) at their terminal ends. Such functionalized fatty acids can have applications in the bio-based chemical industry for the production of value-added chemicals and the replacement of the non-renewable petroleum-based carbon feedstock.

Future perspective – The KASIII that can utilize atypical substrates, such as *A. acidocaldarius*, *C. gingivalis* KASIIIs share structural features with KASIIIs that can process branched chain substrates. However, there may be some specific structural differences that enable binding and catalysis of atypical substrates. Structural studies of KASIIIs identified in chapter 4 will be very helpful in advancing the existing knowledge of KASIII enzymes, and can also help in engineering novel biocatalysts that can utilize substrates with varied groups such as halogens and amines that could enable synthesis of novel bio-based polyesters (e.g.

polyamines, or polymers containing halogen groups for enhanced physiochemical properties).

KASIII can be used for the *in vivo* production of novel ω -functionalized fatty acids

In collaboration with Dr. Huanan Jin, a former post-doctoral fellow in the Nikolau group, we demonstrated that KASIII from *A. acidocaldarius* can process 3-hydroxybutyryl-CoA starter units for the *in vivo* production of ω -1-hydroxy-branched fatty acids in *Rhodospirillum rubrum* (Chapter 4). Production of ω -1-hydroxy-branched fatty acids was made possible via co-expression of a specific KASIII in conjunction with two *R. rubrum* genes, *phaA* and *phaB*, that synthesize 3-hydroxybutyryl-CoA, which is then utilized as a substrate by the recombinant KASIII. *A. acidocaldarius* KASIII (aaKASIII) was used for this proof-of-production of hydroxylated fatty acids because our results showed that aaKASIII is most effective in utilizing hydroxylated substrates (Chapter 4). This result illustrates the practical application of novel KASIII enzymes identified in chapter 4 for the *in vivo* production of unique ω -functionalized fatty acids

Future perspective – The developing technology described in this dissertation is the basis of a start-up company, OmegaChea Biorenewables LLC (see Appendix for the company's business plan), and has vast potential in terms of expanding the fatty acid synthesis metabolic platform to create other diverse fatty acid products, such as fatty acids with amino, sulphate or carboxylic acid groups at their terminal ω -ends. This technology needs further development in the scale-up and production of higher yields of ω -functionalized fatty acids. Re-construction of this metabolic pathway in a lab-friendly organism such as *E. coli* followed by an industrially relevant strain and the future

optimization of production and separations processes can de-risk this technology and make it commercially competent.

Concluding remarks

To summarize, in this dissertation, I have identified residues that are key to KASIII substrate binding and catalysis, however engineering of KASIII to fine-tune its substrate specificity will require either subtle mutations close to the active site or a directed evolution approach to build a functionally robust enzyme. Additionally, I have functionally characterized a dozen previously uncharacterized KASIII enzymes and correlated their diverse substrate specificities with features of their primary structure. This has enabled me to develop a bioinformatics-based structure-function correlation that can be used to predict KASIII functionality, which can be tested quickly using a combination of the *in vivo* and *in vitro* screens described in this dissertation. However, a vast diversity of KASIII enzymes is yet to be explored, and through this work, I have been successful in only touching the tip of the ice-berg. In this work I have provided a detailed characterization of fifteen KASIII enzymes from a collection of approximately 2,300 KASIII genes identified in public databases, which is not even 1% of the existing KASIII diversity. This presents rich opportunities for applying the bioinformatics approach to these KASIIIs to identify candidate enzymes to test further. It is possible that within this large collection of candidate KASIIIs, there exist unique KASIII enzymes with functionalities for aromatic, sulfonated or halogenated precursor molecules, identification of which will diversify products of the fatty acid synthesis pathway.

References

1. Nikolau, B. J., Perera, M. A., Brachova, L., and Shanks, B. (2008) Platform biochemicals for a biorenewable chemical industry. *Plant J* **54**, 536-545
2. Huynh, K. H., Natarajan, S., Song, N. H., Ngo, P. T., Ahn, Y. J., Kim, J. G., Lee, B. M., Eo, Y. D., and Kang, L. W. (2009) Cloning, expression, crystallization and preliminary X-ray crystallographic analysis of beta-ketoacyl-ACP synthase III (FabH) from *Xanthomonas oryzae* pv. *oryzae*. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **65**, 460-462
3. Qiu, X., Janson, C. A., Konstantinidis, A. K., Nwagwu, S., Silverman, C., Smith, W. W., Khandekar, S., Lonsdale, J., and Abdel-Meguid, S. S. (1999) Crystal structure of beta-ketoacyl-acyl carrier protein synthase III. A key condensing enzyme in bacterial fatty acid biosynthesis. *J Biol Chem* **274**, 36465-36471
4. Qiu, X., Janson, C. A., Smith, W. W., Head, M., Lonsdale, J., and Konstantinidis, A. K. (2001) Refined structures of beta-ketoacyl-acyl carrier protein synthase III. *J Mol Biol* **307**, 341-356
5. Davies, C., Heath, R. J., White, S. W., and Rock, C. O. (2000) The 1.8 Å crystal structure and active-site architecture of beta-ketoacyl-acyl carrier protein synthase III (FabH) from *Escherichia coli*. *Structure* **8**, 185-195
6. Scarsdale, J. N., Kazanina, G., He, X., Reynolds, K. A., and Wright, H. T. (2001) Crystal structure of the *Mycobacterium tuberculosis* beta-ketoacyl-acyl carrier protein synthase III. *J Biol Chem* **276**, 20516-20522
7. Musayev, F., Sachdeva, S., Scarsdale, J. N., Reynolds, K. A., and Wright, H. T. (2005) Crystal structure of a substrate complex of *Mycobacterium tuberculosis* beta-ketoacyl-acyl carrier protein synthase III (FabH) with lauroyl-coenzyme A. *J Mol Biol* **346**, 1313-1321
8. Sachdeva, S., Musayev, F. N., Alhamadsheh, M. M., Scarsdale, J. N., Wright, H. T., and Reynolds, K. A. (2008) Separate entrance and exit portals for ligand traffic in *Mycobacterium tuberculosis* FabH. *Chem Biol* **15**, 402-412
9. Rock, C. O., and Cronan, J. E. (1996) *Escherichia coli* as a model for the regulation of dissociable (type II) fatty acid biosynthesis. *Biochim Biophys Acta* **1302**, 1-16
10. Lai, C. Y., and Cronan, J. E. (2003) Beta-ketoacyl-acyl carrier protein synthase III (FabH) is essential for bacterial fatty acid synthesis. *J Biol Chem* **278**, 51494-51503
11. Choi, K. H., Heath, R. J., and Rock, C. O. (2000) beta-ketoacyl-acyl carrier protein synthase III (FabH) is a determining factor in branched-chain fatty acid biosynthesis. *J Bacteriol* **182**, 365-370

12. Heath, R. J., and Rock, C. O. (1996) Inhibition of beta-ketoacyl-acyl carrier protein synthase III (FabH) by acyl-acyl carrier protein in *Escherichia coli*. *J Biol Chem* **271**, 10996-11000
13. Tsay, J. T., Oh, W., Larson, T. J., Jackowski, S., and Rock, C. O. (1992) Isolation and characterization of the beta-ketoacyl-acyl carrier protein synthase III gene (fabH) from *Escherichia coli* K-12. *J Biol Chem* **267**, 6807-6814
14. Choi, K. H., Heath, R. J., and Rock, C. O. (2000) beta-ketoacyl-acyl carrier protein synthase III (FabH) is a determining factor in branched-chain fatty acid biosynthesis. *J Bacteriol* **182**, 365-370
15. Smirnova, N., and Reynolds, K. A. (2001) Branched-chain fatty acid biosynthesis in *Escherichia coli*. *J Ind Microbiol Biotechnol* **27**, 246-251
16. Alhamadsheh, M. M., Musayev, F., Komissarov, A. A., Sachdeva, S., Wright, H. T., Scarsdale, N., Florova, G., and Reynolds, K. A. (2007) Alkyl-CoA disulfides as inhibitors and mechanistic probes for FabH enzymes. *Chem Biol* **14**, 513-524
17. Tsay, J. T., Oh, W., Larson, T. J., Jackowski, S., and Rock, C. O. (1992) Isolation and characterization of the beta-ketoacyl-acyl carrier protein synthase III gene (fabH) from *Escherichia coli* K-12. *J Biol Chem* **267**, 6807-6814
18. Jackowski, S., and Rock, C. O. (1987) Acetoacetyl-acyl carrier protein synthase, a potential regulator of fatty acid biosynthesis in bacteria. *J Biol Chem* **262**, 7927-7931
19. Jackowski, S., Murphy, C. M., Cronan, J. E., and Rock, C. O. (1989) Acetoacetyl-acyl carrier protein synthase. A target for the antibiotic thiolactomycin. *J Biol Chem* **264**, 7624-7629
20. Qiu, X., Choudhry, A. E., Janson, C. A., Grooms, M., Daines, R. A., Lonsdale, J. T., and Khandekar, S. S. (2005) Crystal structure and substrate specificity of the beta-ketoacyl-acyl carrier protein synthase III (FabH) from *Staphylococcus aureus*. *Protein Sci* **14**, 2087-2094
21. Choi, K. H., Kremer, L., Besra, G. S., and Rock, C. O. (2000) Identification and substrate specificity of beta -ketoacyl (acyl carrier protein) synthase III (mtFabH) from *Mycobacterium tuberculosis*. *J Biol Chem* **275**, 28201-28207
22. Pereira, J. H., Goh, E. B., Keasling, J. D., Beller, H. R., and Adams, P. D. (2012) Structure of FabH and factors affecting the distribution of branched fatty acids in *Micrococcus luteus*. *Acta Crystallogr D Biol Crystallogr* **68**, 1320-1328
23. Singh, A. K., Zhang, Y. M., Zhu, K., Subramanian, C., Li, Z., Jayaswal, R. K., Gatto, C., Rock, C. O., and Wilkinson, B. J. (2009) FabH selectivity for anteiso branched-chain fatty acid precursors in low-temperature adaptation in *Listeria monocytogenes*. *FEMS Microbiol Lett* **301**, 188-192

24. Khandekar, S. S., Gentry, D. R., Van Aller, G. S., Warren, P., Xiang, H., Silverman, C., Doyle, M. L., Chambers, P. A., Konstantinidis, A. K., Brandt, M., Daines, R. A., and Lonsdale, J. T. (2001) Identification, substrate specificity, and inhibition of the *Streptococcus pneumoniae* beta-ketoacyl-acyl carrier protein synthase III (FabH). *J Biol Chem* **276**, 30024-30030
25. Han, L., Lobo, S., and Reynolds, K. A. (1998) Characterization of beta-ketoacyl-acyl carrier protein synthase III from *Streptomyces glaucescens* and its role in initiation of fatty acid biosynthesis. *J Bacteriol* **180**, 4481-4486
26. Wen, L., Chmielowski, J. N., Bohn, K. C., Huang, J. K., Timsina, Y. N., Kodali, P., and Pathak, A. K. (2009) Functional expression of *Francisella tularensis* FabH and FabI, potential antibacterial targets. *Protein Expr Purif* **65**, 83-91
27. Gajiwala, K. S., Margosiak, S., Lu, J., Cortez, J., Su, Y., Nie, Z., and Appelt, K. (2009) Crystal structures of bacterial FabH suggest a molecular basis for the substrate specificity of the enzyme. *FEBS Lett* **583**, 2939-2946
28. Brown, A. K., Sridharan, S., Kremer, L., Lindenberg, S., Dover, L. G., Sacchettini, J. C., and Besra, G. S. (2005) Probing the mechanism of the *Mycobacterium tuberculosis* beta-ketoacyl-acyl carrier protein synthase III mtFabH: factors influencing catalysis and substrate specificity. *J Biol Chem* **280**, 32539-32547
29. Kaneda, T. (1991) Iso- and anteiso-fatty acids in bacteria: biosynthesis, function, and taxonomic significance. *Microbiol Rev* **55**, 288-302

APPENDIX

BUSINESS PLAN FOR OMEGACHEA BIORENEWABLES LLC

Executive Summary

Mission

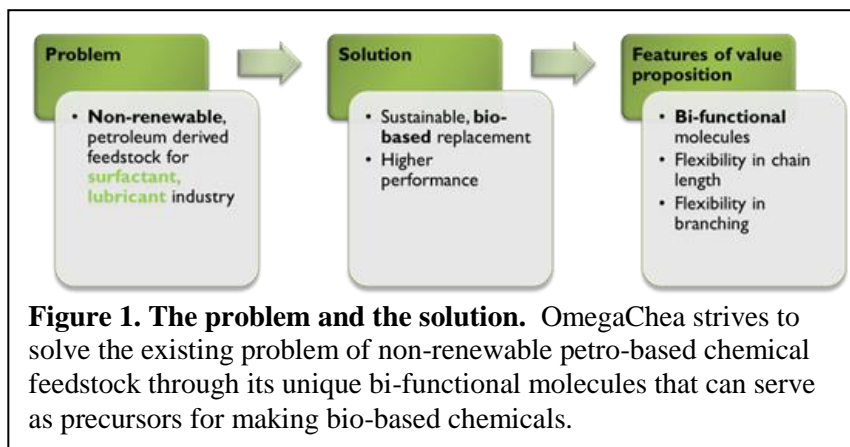
OmegaChea Biorenewables LLC is a startup company that will commercialize a new scalable manufacturing capability by designing microorganisms to produce bio-based chemicals for surfactants, lubricants and polymers.

Background

The core technology of OmegaChea was developed based on the KASIII technology described in chapter 3 of this dissertation as a part of the NSF-funded Engineering Research Center for Biorenewable Chemicals (CBiRC) at Iowa State University. The core business concept and commercial feasibility was explored and developed within an NSF I-Corps award made to the OmegaChea founders.

Company Status

OmegaChea is a start-up company and is in the initial stages of research and development.



OmegaChea has received an i6 Green grant from the State of Iowa to develop its technology, and has secured \$225,000 from NSF STTR Phase I award in July 2013. It aims to apply for

Phase II of NSF STTR award in Dec 2014. Status of the OmegaChea technology is described separately in a section below.

The Problem

The vast majority of the current feedstock for the chemical industry is petroleum derived, and is imported to a large extent. Moreover, petroleum-based chemical precursors do not offer flexibility in their structures.

The Solution

Fatty acids provide attractive bio-based alternative “green” chemicals for both biopolymers and surfactants or lubricants. They are attractive because they offer chemical structures and physical properties that are similar or even better than the petroleum-derived chemicals that are commonly used by these industries today. OmegaChea will produce bi-functional fatty acids that can replace current petroleum-based feedstock chemicals (Fig. 1).

Product and the Technology

The technological basis for OmegaChea is the bioengineering of a key enzyme that determines the chemical nature of the ω -end (the omega-end) of fatty acid molecules. OmegaChea will use diverse versions of this enzyme that exert an exquisite degree of control over the synthesis of ω -functionalized fatty acids. OmegaChea’s objective is to commercialize the use of these enzymes in a bioengineered microbial manufacturing platform to sustainably produce bi-functional chemicals.

Market Opportunity

The OmegaChea technology will address two potential market opportunities with new bio-based chemicals. The first of these is the polymer industry (plastics of many types),

which currently uses bi-functional monomers (i.e., a molecule with chemical functionalities at both ends) that are predominantly generated from non-sustainable petroleum feedstocks. The second market opportunity is the surfactant/lubricant industry, which is seeking bio-based bi-functional fatty acids (specifically branched chain fatty acids) that offer enhanced chemical-physical lubricity and tribological properties. The surfactant / lubricant industry has lower barriers to entry and therefore will be the market initially targeted.

Business Model

The core business concept of OmegaChea was developed by the CBiRC team within an NSF I-Corps award (NSF award #IIP-1237247). The commercial potential of this technology was explored in conversations with possible OmegaChea customers, which include Genencor, BASF, DSM, Procter & Gamble, and Cargill Industrial Oils and Lubricants. These conversations identified an initial customer-base, and led to development of a business model detailed in the business plan.

Typical Customers

Through the NSF I-Corps program, an initial base of customers consisting of surfactant and lubricant formulators/manufacturers was identified for OmegaChea. Procter and Gamble, which said, *“You give us 1kg of product, we can evaluate”*, will be one of our early-vangelists. In addition to P&G, several other surfactant, lubricant manufacturers expressed interest in our technology.

Minimum Viable Product

Although the patent-pending OmegaChea technology has the potential to produce a variety of different ω -functionalized fatty acids (i.e., hydroxy-, halo-, branched-, or cyclic-

fatty acids), OmegaChea will initially target the production of ω -hydroxy-branched fatty acids, which have applications in both surfactants (detergents, household cleaners) and lubricants (motor lubricants, cutting tool lubricants) and in polymers (i.e., polyesters).

Status of the Technology

OmegaChea has been successful in providing a proof-of-principle of production of bi-functional fatty acids, specifically ω -hydroxy-branched fatty acids in engineered microbial hosts, and has applied for a patent for this technology. The next step for OmegaChea is to optimize the production conditions and then scale-up the yields of the bi-functional fatty acids, and overcome any toxic effects of these products on microbial hosts.

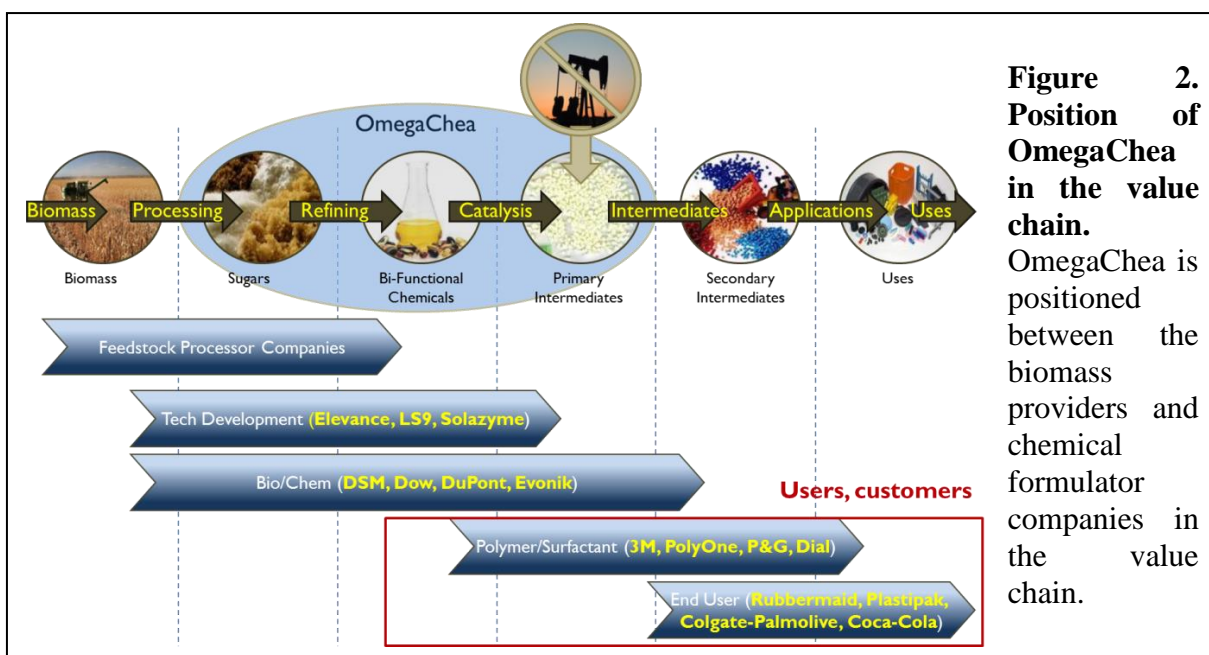
Financial Projections

For proof-of-concept development, OmegaChea will require upto \$250,000 over next one year. For prototype development, it will need additional \$750,000 for a period of 2-3 years. Following prototype development, we will seek funding from venture capitalists (up to \$5million) for pre-pilot scale and then pilot scale development. For scale-up, we might partner with third-party scale up providers (like Tate and Lyle, Poet) to reduce our capital investment costs.

We envision that we will generate systems that produce products at different scales of yield, and our technology will enable the replacement of existing petroleum-based molecules with unique, bio-based bi-functional molecules.

Background

OmegaChea Biorenewables LLC is a startup entity that is seeking to expand upon and commercialize technological innovations made by the NSF-funded Engineering Research Center for Biorenewable Chemicals (CBiRC). This technology has the potential to impact the development of the local and national objective of growing the bioeconomy using advanced manufacturing processes. OmegaChea's initial target is to develop a new “designer” microorganism that will become a key part of a scalable bio-based process that manufactures bi-functional fatty acids, and subsequently establish a new sustainable bio-manufacturing platform for chemicals. Multiple biorenewable products can be envisioned that will have



customers in the polymer, lubricant and surfactant industries. A dialog has already been established between OmegaChea and some existing commercial entities that currently derive their chemical intermediates from petroleum sources, but see opportunities and premiums to switching to bio-based feedstocks. Therefore, OmegaChea will provide these customers access to new “designer” molecules that will represent an alternative, sustainable resource for

their chemical needs (Fig. 2). OmegaChea expects to reach the marketplace through close interactions and partnering with larger companies. This partnering relationship is being developed from existing exposure and interactions with the significant member company portfolio of CBiRC.

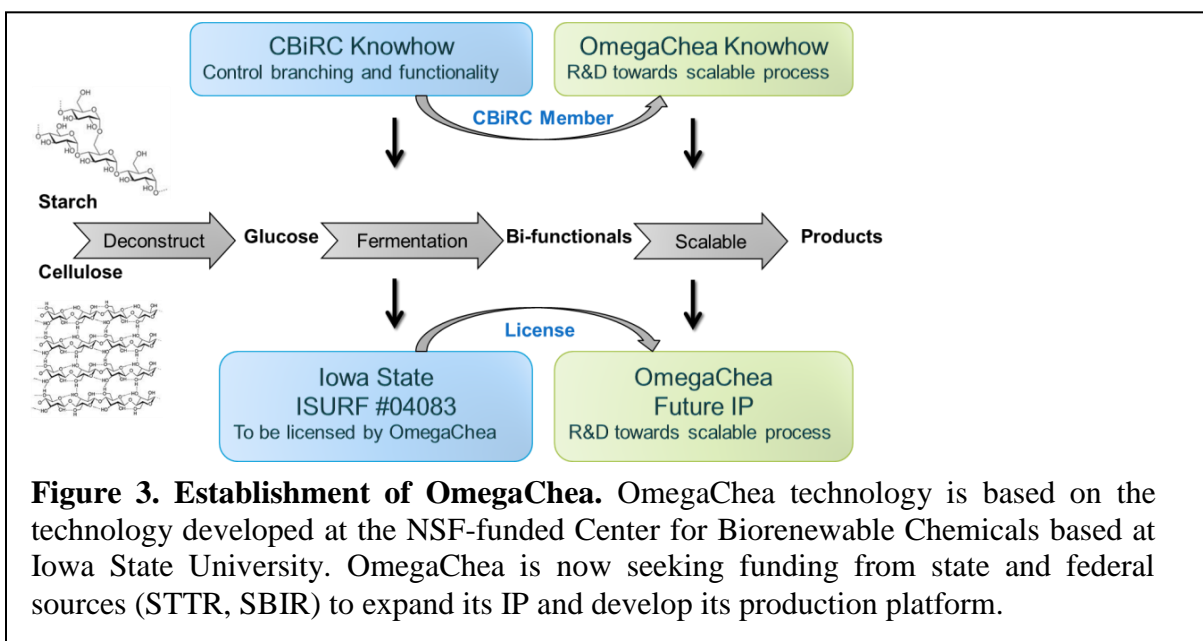
Founding History

The concept of OmegaChea began to take shape for the Founders in December 2011, and as a startup entity the company is at an early stage of development. The technology that OmegaChea is seeking to commercialize has its origins in the PhD research studies of Ms Shivani Garg, which was supported in part by CBiRC funding (Fig. 3). Co-founder Dr. Basil Nikolau serves as Ms Garg's PhD major professor, and Co-founder Dr. Yandeau-Nelson is Ms Garg's co-mentor in Iowa State University's Molecular, Cellular and Developmental Biology PhD program.

Technology-led entrepreneurship course and the NSF I-Corps program

As the Founding Interim President of OmegaChea, Ms Garg has established a startup commercial entity, which provides her with an opportunity to expand her technical expertise, and integrate her entrepreneurial drive. Ms Garg has taken multiple initiatives to expand her expertise in entrepreneurship. Initially this was via a formal course at ISU; "Technology-led Entrepreneurship" taught in Spring 2011 by Dr. Peter Keeling, Innovation Director of CBiRC and co-founder of OmegaChea. Subsequently, as the concept of OmegaChea as a commercial entity began to take shape, Garg, Nikolau and Keeling began exploring the commercial potential of the startup with funds from an NSF I-Corps award (award #IIP-1237247). In the I-Corps program, the OmegaChea team (Garg as the Entrepreneurial Lead,

Nikolau as the Principal Investigator, and Keeling as the Industry Mentor) explored the potential customer base for the OmegaChea technology.



Management Team

The OmegaChea management team consists of the four Founders of OmegaChea (Garg, Nikolau, Yandeau-Nelson and Keeling) who developed the technological and business aspects of the start-up. Their expertise and roles within this project are described below.

Shivani Garg (Founder and President, OmegaChea)

Garg is currently a graduate student conducting research towards a PhD in Molecular, Cellular and Developmental Biology with a co-major in Biorenewables at Iowa State University. She obtained her Bachelor of Engineering at Panjab University, India majoring in Biotechnology Engineering (2002-06), and then worked with Infosys, India a leading software development company (2006-08). Ms Garg joined the Nikolau group in 2009 for

her graduate studies, and began the CBiRC-sponsored project on KASIII technology that is at the core of the OmegaChea innovation. As the entrepreneurial lead on the NSF I-Corps project, she explored the commercial potential of KASIII technology and developed a business model, which is the basis for foundation of OmegaChea as a business entity. She is also the co-PI on an i6 grant awarded by State of Iowa to further advance the OmegaChea technology towards commercialization.

Marna Yandea-Nelson (Founder and Vice-President Research Operations, OmegaChea)

Dr. Yandea-Nelson is trained in plant molecular genetics and molecular biology (Ph.D. 2005) and biochemistry (post-doc, Penn State). Since 2009 Yandea-Nelson (Associate Scientist and Graduate Faculty) has coordinated research efforts within CBiRC. Yandea-Nelson's role as Interim Vice-President Research Operations in this STTR is to directly manage the proposed project, and particularly to liaise and coordinate the management of the efforts between OmegaChea and the CBiRC researchers.

Basil J. Nikolau (Founder and Chief Scientific Officer, OmegaChea)

Dr. Nikolau is the Frances M. Craig Professor of Biochemistry at Iowa State University. Dr. Nikolau is trained in the areas of biochemistry and molecular biology. He has published nearly 100 peer-reviewed manuscripts in the field, being supported by research grants from the NSF, the USDA, US DoE, and the NIH. Currently he serves as the Deputy Director of the NSF-funded Engineering Research Center, CBiRC, based at Iowa State University. Nikolau's role as the Interim Chief Scientific Officer of OmegaChea is to provide guidance and expertise to the collaboration between OmegaChea and the ISU research of Ms Shivani Garg, to ensure the technical success of the combined project.

Peter L. Keeling (Founder and Vice-President Business Operations, OmegaChea)

Dr. Keeling is trained in biochemistry and has more than twenty-five years of industry experience in the biotechnology sector. Dr. Keeling's current position is Director of Innovation with the NSF-funded CBiRC. Dr Keeling is Founder and Director of startup companies (ExSeed 1994; EnaGen 2007; EneGea 2009; GlucanBio 2011) that have focused on technology development. In his role within CBiRC, his efforts have resulted in forming the CBiRC's Biobased Foundry. Dr Keeling also coordinates with ISURF and the State of Iowa Economic Development Authority to steer investment dollars into startups. Since the inception of CBiRC's Biobased Foundry, this effort has nurtured 6 startup entities, including GlucanBio, SusTerea, SolysTE, RecyclaR, Abios as well as the current startup, OmegaChea. Keeling's role as the Interim Vice-President for OmegaChea is to provide guidance and expertise to the collaboration between OmegaChea and the CBiRC and steer the commercialization and investment strategy for the growth and success of OmegaChea.

Problem and its solution**Problem**

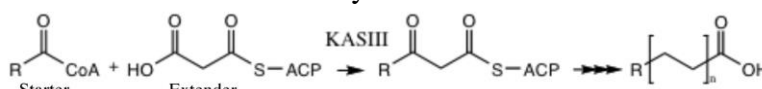
The US imports almost 10 million barrels of petroleum a day (1), to create a multi-billion dollar plastics and specialty chemical industry that obtains its monomers from petroleum feedstocks. Currently, only a limited number of bio-based products are available in the market including polylactic acid (PLA), polyhydroxy-butyrates (PHB), polyethylene terephthalate (PET) based on 1,3-propanediol, and emerging products based on succinic acid and adipic acid (2). However, the global marketplace for the bio-plastics "green" market is

projected to expand to over a billion dollars (3). With increasing awareness of reduced environmental impacts of bio-based plastics, the market for these products will continue to grow.

Solution

In this project, OmegaChea will develop a new, proprietary microbial fermentation host (a unique *E. coli* strain) that will produce scalable quantities of ω -hydroxy-branched-fatty acids. ω -Hydroxy-branched-fatty acids rarely occur in nature, but they offer many new opportunities for the chemical industry, particularly with polymer, lubricant and surfactant applications. *E. coli* has been chosen for this as a manufacturing host because

Reaction 1. KASIII-mediated synthesis of bi-functional fatty acids.



it offers many technological advantages, and it is widely used in industrial fermentation capabilities.

The proposed innovation is based on new gene technology developed by CBiRC researchers (Garg, Yandea-Nelson and Nikolau), which includes identification and characterization of diverse 3-ketoacyl-ACP synthase III (KASIII) enzymes. KASIII catalyzes the initial reaction of fatty acid biosynthesis using an acyl-CoA primer as its substrate (Reaction 1). By bioengineering a unique *E. coli* organism with novel KASIII enzymes that can use unusual substrates like hydroxy-acyl-CoA and/or branched-acyl-CoA primer-substrates, it is OmegaChea's proposition that a new microbial based manufacturing platform can be designed to produce novel ω -hydroxy-branched-fatty acids that are not known to occur in nature, but have wide-ranging applications in chemical industries.

Product Details

Specifically, our technology will enable production of bi-functional, bio-based fatty acid products. We are initially targeting the production of hydroxy-branched-fatty acids, based on market feedback and following reasons –

1. These are bi-functional carboxylic acids and serve as excellent handles for chemical reactions.
2. Chain lengths of these acids can be customized between C8-C16 and offers flexibility to detergent, lubricant manufacturers.
3. Positions and number of branches in these molecules can also be customized, which improves their performance as detergents and lubricants at lower temperatures by preventing solidification of these molecules at low temperatures.

These molecules can be used as chemical feedstocks in the manufacture of detergents, surfactants, lubricants and specialty chemical synthesis. Therefore, typical customers of the technology would be specialty chemical producers and surfactant/lubricant manufacturers, like companies such as Procter and Gamble, DSM, BASF, and Evonik. Additionally, biotechnology companies could be possible customers, but such companies may be more interested in the genetic elements that encode our bioengineered enzymes or bacterial strains that express these enzymes.

Status of technology and critical milestones

OmegaChea has demonstrated proof-of-principle by producing ω -hydroxy-branched-fatty acids in an engineered microbial host (*Rhodospirillum rubrum*) which has novel KASIII genes in it. The next key milestone for OmegaChea is to engineer the pathway for production of ω -hydroxy-branched-fatty acids in a lab friendly microbe, i.e. *Escherichia coli* and then

optimize the production process. Currently, we are working on optimization of the KASIII biocatalysts that are at the core of this technology. After biocatalyst optimization, next step will be *E. coli* platform construction and optimization, followed by efforts to scale-up the production of bi-functional fatty acids (Fig. 4).

To get the first viable product to marketplace, OmegaChea must achieve certain critical milestones, which include:

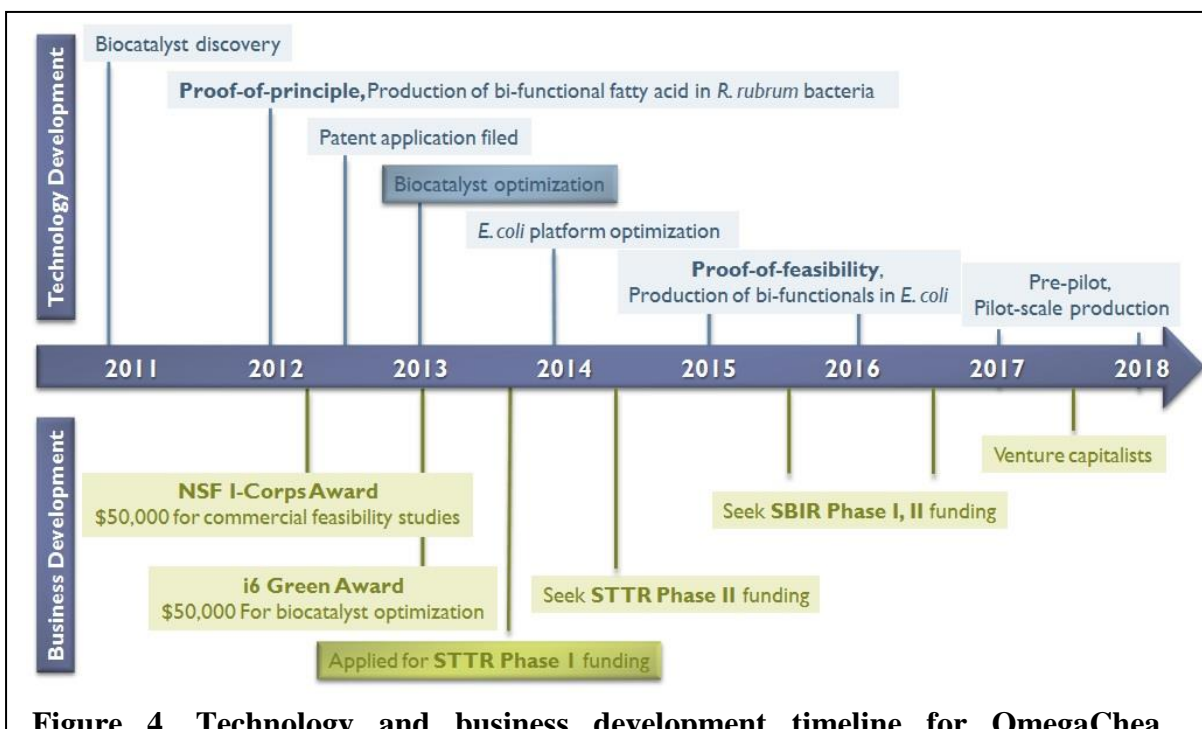


Figure 4. Technology and business development timeline for OmegaChea. Currently, OmegaChea is in the biocatalyst optimization stage, and has funding from i6 Green grant. The next goal for OmegaChea will be to construct *E. coli* platform using optimized biocatalysts to produce bi-functional fatty acids in *E. coli*. OmegaChea has recently been awarded \$225,000 NSF STTR Phase I funding to support its research and development plans.

- i) Demonstration of proof-of-concept of production of bi-functional fatty acids in *E. coli* and other microbial hosts.
- ii) Identification of optimized KASIII biocatalysts for the maximal production of hydroxy-branched-fatty acids.

- iii) Yield optimization of bi-functional fatty acids by engineering biocatalysts in appropriate microbial hosts.
- iv) Optimization to minimize the processing and separations costs.
- v) Scale-up by partnering with third-party scale-up providers; and product testing by partnering with earlyvangelists.
- vi) Finally, product customization and initial market launch.

Competitive Advantage

The value proposition for the product

The value proposition of OmegaChea is to translate early-stage technology into a viable, de-risked and scalable process to manufacture, use and sell bio-based products. If it is successful, OmegaChea will have developed a portfolio of technologies that deliver an array of marketable bi-functional chemicals. In this form the company will become a target for acquisition by one of its partnering large multinationals. More specifically OmegaChea will have a competitive advantage in market because of the following value proposition for its products -

- i) Sustainable raw materials for surfactant, lubricant industry
- ii) Flexibility in molecular carbon chain-length and branching of chemical products
- iii) Higher performance surfactants and lubricants due to molecular branching of chemical products
- iv) Green 'fermentative' technology

This value proposition will be offered by our technological advances that will spur new bio-based manufacturing platforms for bi-functional carboxylic acids (hydroxy-fatty acids, branched chain fatty acids) with chain lengths that can be customized to between 8 and 16

CONFIDENTIAL

Table 1. Product features, value proposition, applications and advantages over competitor's products.

Minimum Viable Product	Customer Segments	Specifications and advantages over competitor's products
ω-Hydroxy-Branched-Fatty Acid	Surfactants	<ul style="list-style-type: none"> • Branching improves surfactant performance at colder temperature • Position of branches and chain length can be customized
	Lubricants	<ul style="list-style-type: none"> • Branching improves lubrication properties at colder temperature • Position of branches and chain length can be customized
	Polymers	<ul style="list-style-type: none"> • Hydroxyl group makes a bi-functional fatty acid that can be polymerized to give bio-based polymers • Chain length can be customized

carbon atoms. A summary of minimum viable products of OmegaChea for each market segment and their specifications are provided in Table 1.

Green technology with better performance

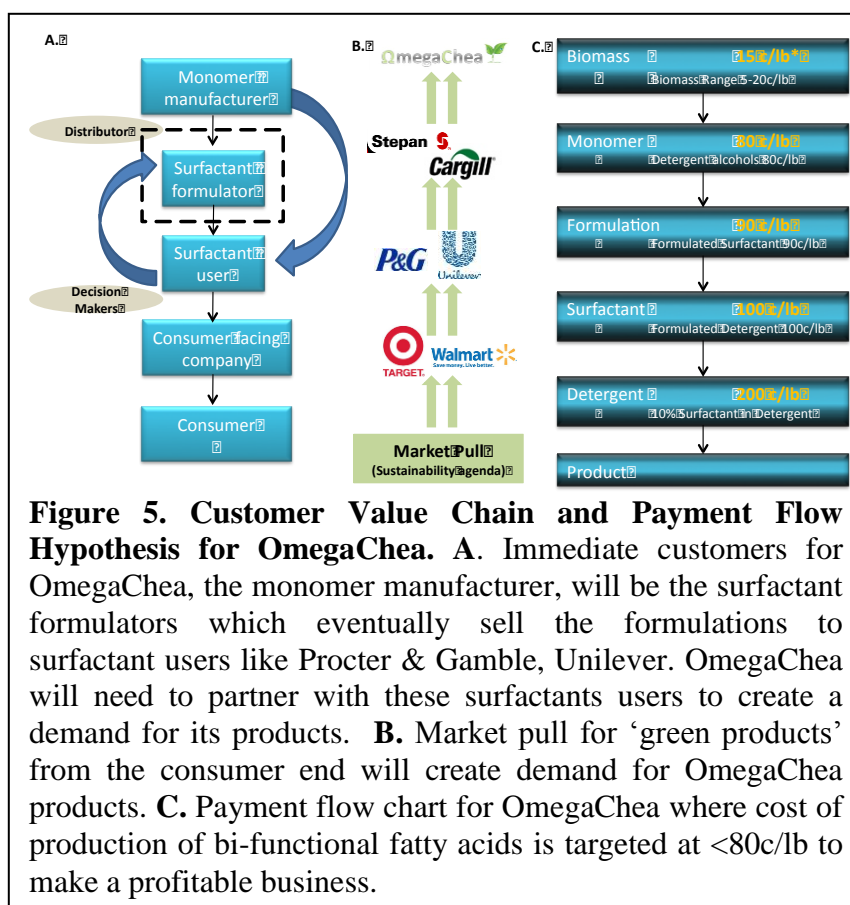
The value for customers lies in a 'green' product that offers better performance than the existing petroleum based chemical precursor molecules. Some of the customers we talked to during the I-Corps program indicated that they will be willing to pay a premium for a bio-based product as long as it has the same or better performance than the petroleum derived products. Additionally, some companies, like Procter and Gamble, BASF, DuPont reserve a portion of their budget for developing bio-based technologies. Such chemical companies would be our first target customers.

Competition

Three potential market segments exist for our bi-functional fatty acid products. Each of the market segments offers a different competitive landscape, with the polymers market

being fiercely competitive, offering very limited room for a new bio-based startup, like OmegaChea to find its own niche. In contrast, the surfactant and lubricant markets are relatively less crowded with bio-based products and hence, they provide more opportunities for OmegaChea to grow.

The main competition for OmegaChea is from other early-stage bio-based chemicals technology developers such as Amyris and Metabolix. These companies have similar aspirations to exploit the microbial enzymatic machinery to produce a wide-array of bio-based



chemicals and fuels. Some of these bio-based companies have partnered with giant chemical manufacturers (like Unilever, Procter & Gamble). These partnerships are designed to create a consistent and long-term demand for their products. OmegaChea will adopt a similar strategy in that, after initial technological de-risking, it will partner with its early-vangelists identified through the I-Corps program, and seek synergies within CBiRC’s Innovation Ecosystem

(e.g., DSM, Evonik and Procter & Gamble) to create a supply-demand chain for its products and begin to establish itself in the marketplace.

To be competitive in market, OmegaChea will have to bring its production costs under 80c/lb and sell its products at 80c/lb to make profits. Market pull for the bio-based products will help in competing against established petroleum-based competitors' products. Additionally, strategic partnership with giant companies like P&G, Unilever will help create demand for OmegaChea's bio-based products (Fig. 5).

Customer Segments

By providing access to new bio-based chemical feedstocks, the OmegaChea technology will address the needs of three potential customer segments –

- i) Surfactants formulators,
- ii) Lubricants formulators, and
- iii) The bio-based polymer manufacturers.

The surfactant and lubricant formulators are seeking bio-based branched-chain fatty acids that offer enhanced chemical-physical lubricity and tribological properties, particularly at low temperature application. The polymer industry uses bi-functional monomers that can be used to chemically react and polymerize to form, for example polyesters or polyamides. All of these customer segments currently derive their feedstocks predominantly from non-sustainable petroleum sources. Modified fatty acids that will be offered by OmegaChea provide attractive bio-based alternative “green” chemicals for both these applications. This is because they offer chemical structures and physical properties that are similar or even

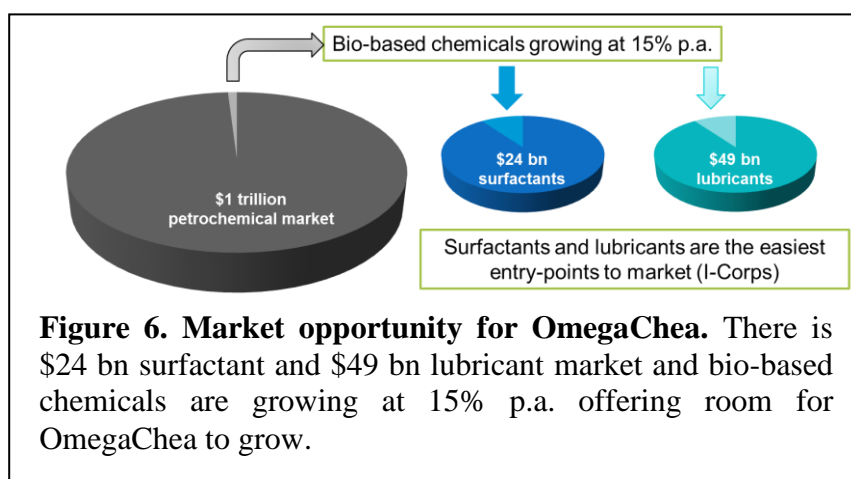
better than the petroleum-derived chemicals that are commonly used by these industries today.

Market Opportunity

The market opportunity for OmegaChea was explored and defined by market studies and meetings with industry representatives that occurred within the NSF-funded I-Corps grant. These studies and interactions with CBiRC member companies revealed that there are

4-5 different products that could be developed with the new KASIII-manufacturing process.

OmegaChea's first viable product will aim to target the surfactants



market segment that has significant value to potential customers. Eventually, OmegaChea's bio-based market segments will include surfactants and lubricants formulators and the bio-based polymer manufacturers. The surfactants and lubricants markets have a customer-driven premium on the bi-functional chemicals that can be impacted by OmegaChea's technology. The lubricant and surfactant formulators desire ω -branched fatty acids, which provide added functionality in low temperature lubrication applications. In contrast, the bio-based polymer manufacturers desire ω -hydroxy-fatty acids, which can be used in the assembly of polyesters, a process that is much less transparent to the consumer. Thus, the initial product that OmegaChea has targeted for manufacture meets all of these specifications by incorporating hydroxyl-group functionality in the ω -1 position of fatty acids, providing both a branched-

structure for improved lubrication and surfactant properties, whereas the hydroxyl-group can be used in other reactions to assemble polyesters as is typically done in chemical industry.

Typical Customers

Some examples of our typical customers in each of the three segments are Procter & Gamble (www.pg.com), Seventh Generation (www.seventhgeneration.com), and Target Up & Up (www.target.com) in the surfactants segment; Cargill Industrial Oils and Lubricants (www.cargill.com), and Stepan (www.stepan.com) in the lubricants segment; and BASF (www.basf.com), DSM (www.dsm.com), and Danisco/DuPont (www.dupont.com) in the bio-based polymers segment. We also identified some earlyvangelists for our products through the I-Corps program, namely Procter & Gamble and Evonik (www.evonik.com), which will be our first target customers as we move closer to commercializing our product.

Market Size and Market Share

The total market size for petrochemicals industry is estimated to be \$1 trillion, of which bio-based chemicals share a small percentage. However, market for bio-based chemicals grows at 15% per annum. The main market share for OmegaChea will be the growing 15% bio-based sector of the \$24 billion surfactants and \$49 billion lubricants markets (Fig. 6).

Distribution Channels

OmegaChea will use the indirect channel of wholesale distributors for distributing its products. During the I-Corps program, OmegaChea identified various distributor networks operating in mid-west that distribute bio-based chemicals similar to OmegaChea's first target

products. National Association of Chemical Distributors (NACD) will be another channel that will help OmegaChea reach its customers throughout the country.

Entry Barriers

Our I-Corps sponsored discussions with representatives of the polymer industry indicate that the polymer segment is a broad market and OmegaChea would find it difficult to enter this market as a startup entity. The main reason is that polymer manufacturers are large, integrated entities that have their own capital investments to synthesize feedstock monomers. Therefore, it is challenging for a bio-based startup company (such as OmegaChea) to dislodge the existing petroleum based feedstock system in the polymer markets. On the other hand, the surfactants and lubricants markets, which are each worth more than \$20 billion, have separate entities for formulation and utilization. Entry barriers are much lower than the polymers industry, and wide-ranging applications are possible. Additionally, there is a strong market pull from the consumer end as demand increases for green, sustainable products (Fig. 5). This market pull helps shift focus of the surfactant users/formulators towards feedstocks that are bio-based. Therefore, OmegaChea's initial market-target is a product that has potential in both the polymer and surfactant industries, but because of the findings of our I-Corps investigations the priority for OmegaChea will be the surfactant formulators.

Marketing Plan

OmegaChea's marketing strategy will involve co-creating value proposition with the customer to earn customer trust and create customer satisfaction and good-will. Active participation in bio-based products/technology conferences, such as ICIS World Surfactant Conference, International Conference on Green and Sustainable Chemistry, and trade-shows

will help OmegaChea reach its target customer base effectively. Joining customer communities and surveying their needs/changing demands will help us tailor our marketing plan as OmegaChea evolves (Table 2).

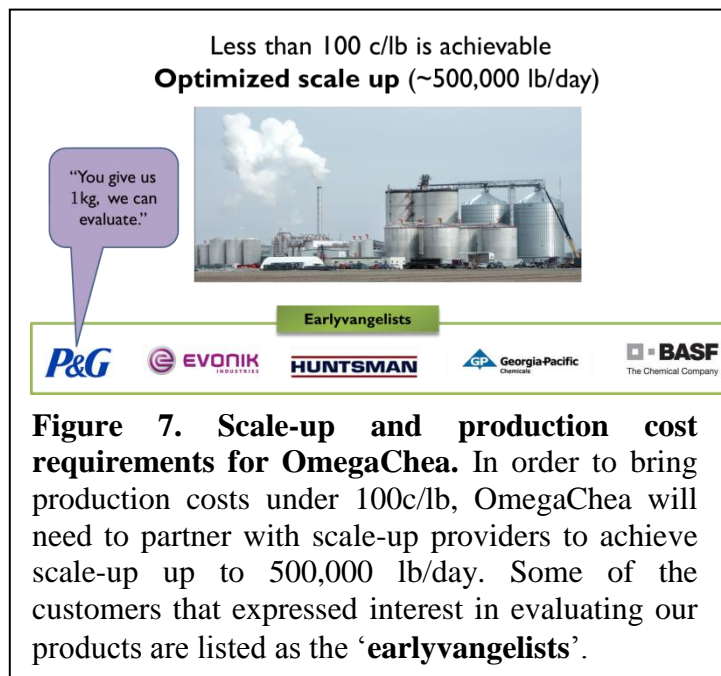
Table 2. Marketing Plan for OmegaChea

Task	Strategy	% of Marketing Budget
Spread awareness	1. Trademarks 2. Web site with ordering capability 3. Measurable metrics	20%
Get customers • Connect	1. Press releases 2. Trade-shows 3. ICIS surfactant conference 4. American Oil Chemists Soc. Meeting 5. LinkedIn	40%
Keep customers • Innovate • Partner	1. Customer communities 2. Co-create value proposition 3. Customer surveys 4. Promotional offers	20%
Grow customers	1. Referrals program 2. Trade-shows, conferences	20%

Operations Plan

The key challenge for OmegaChea will be to manufacture and sell its products at a price that is competitive with the current market prices of surfactants and lubricants. Based on our discussions with various bio-based industry representatives and our techno-economic analysis, we have discovered that to build a profitable entity, the target cost of production for bi-functional fatty acids should be less than 100¢/lb (Fig. 7). This cost was calculated by considering current biomass costs (15¢/lb) and the current estimated costs of surfactants (100¢/lb). The eventual customers will be formulation companies that typically pay 90-100¢/lb for raw monomers, which are formulated, and then sold to surfactant and lubricant manufacturers. To meet our target costs of <100¢/lb for bi-functional monomers, we will ultimately need a highly scalable process. This requires three mega-deliverables:

- i) Pre-pilot-scale optimization of the fermentation and separations processes;
- ii) Scale up to pilot-scale production that demonstrates the scalability of the overall process, and
- iii) Commercial-scale production will require at least ~500,000 lb/day, as is



done by other value-added bio-based chemical manufacturers.

Financial Projections and Plan

Until OmegaChea reaches pilot-scale production stage, we envision seeking additional resources from State (e.g., State of Iowa Development Fund and the ISURF Technology Funds for Applied Research) and federal funding sources (e.g., SBIR, STTR Phase II and PFI-AIR). Because OmegaChea is a CBiRC-associated startup, the STTR and NSF-PFI-AIR funding programs will be specifically targeted, as these are programs designed for the development of innovations from ERCs. The envisioned future federal funding will help foster partnerships with larger companies identified through the I-Corps program, and CBiRC’s Innovation Ecosystem. We envision working closely with these larger entities to explore partnering to fund the scaling-up of our technology to a level needed to compete in the marketplace. There are significant financial challenges ahead for the expected commercialization. Even pilot-scale production may require investments in the order of a few

million dollars and full-scale production will require even greater investments (Fig. 8). These kinds of investments are hard to reach by a small entity, making it necessary to envision acquisition rather than doing it alone. OmegaChea anticipates that the unique exposure to CBIIRC's diverse innovation ecosystem, combined with partnering arrangements with larger entities will reap a long-term benefit.

	2013	2014	2015	2016	2017	2018
R&D DELIVERABLES						
<i>E. coli</i> based platform for manufacturing ω -functionalized fatty acids (STTR)						
Prototype development of the fermentation and separations systems						
Pre-pilot-scale optimization of the fermentation and separations process						
Scale up to pilot-scale production demonstrating scalability						
Commercial-scale production (requires at least ~500,000 lb/day)						
BUSINESS DELIVERABLES						
Company Formation and Preliminary Business Strategy Development						
Intellectual Property Development						
Pilot Scale Company Partnerships for Commercial Development						
REVENUE PROJECTIONS (Sources of Funding)						
I6-Green (Characterization of biocatalysts)	\$100k					
State of Iowa, Development Funds (Biocatalyst Optimization)	\$50k	\$100k				
STTR Phase-I (<i>E. coli</i> platform proof of concept)	\$100k	\$125k				
STTR Phase-II (Prototype development)		\$150k	\$400k	\$200k		
SBIR Phase-I (Pre-pilot scale proof of feasibility)			\$150k			
SBIR Phase-II (Pre-pilot scale development)				\$150k	\$400k	\$200k
Company Partnerships (Pre-Commercial Pilot Scale Development)						\$5,000k
TOTAL	\$250k	\$375k	\$550k	\$350k	\$400k	\$5,200k

Figure 8. Revenue projections for OmegaChea for next 5 years to meet its R&D and business deliverables (CONFIDENTIAL)

References

1. U. S. E. I. Administration (2011) Monthly Energy Review, http://www.eia.gov/energy_in_brief/foreign_oil_dependence.cfm.
2. Frost and Sullivan (2009) Global bio-based plastics market, <http://www.frost.com/prod/servlet/report-toc.pag?repid=M4A1-01-00-00-00>.
3. Ceresana Research (2011) Market Study: Bioplastics, <http://www.ceresana.com/en/market-studies/plastics/bioplastics>.